Remarks

The Amendments

Claim 1 has been amended to delete the recitations "or a functional equivalent thereof" and "solely." Claim 1 also has been amended to recite that "the heterologous gene expresses an active pyruvate decarboxylase." This amendment is supported by Example 10 of the specification, which discloses expression of an active pyruvate decarboxylase from a heterologous gene.

Claims 9, 12, and 13 have been rewritten as independent claims.

Claim 14 has been amended to recite bacterial strain "LN-DP1" in place of "LN-P1." The amendment corrects an obvious typographical error. See Example 6, which describes the "Production of Strain LN-DP1." (Page 10, line 22 to page 11, line 5.)

New claims 26-28 recite that the bacterium of claims 9, 12, and 13 is a thermophile. These new claims are supported by originally filed claim 3.

New claims 29-31 recite that the bacterium of claims 9, 12, and 13 further comprises an inactivated lactate dehydrogenase gene. New claims 29-31 are supported by originally filed claim 5.

None of these amendments introduces new matter or requires a new search.

Rejection of claims 1-13 under 35 U.S.C. §112, second paragraph

Claims 1-13 are rejected under 35 U.S.C. §112, second paragraph as being indefinite. Applicants respectfully traverse.

The Office Action asserts that the recitation "but has solely native alcohol dehydrogenase function" in claim 1 is indefinite. As suggested in the Office Action, this recitation has been replaced by the recitation "has native alcohol dehydrogenase function."

The Office Action also asserts that the recitation "or a functional equivalent thereof" in claim 1 is indefinite. To advance prosecution, this recitation has been deleted.

Withdrawal of these rejections of claims 1-13 is respectfully requested.

Rejection of claims 1-14 under 35 U.S.C. §112, first paragraph

Claims 1-14 are rejected under 35 U.S.C. §112, first paragraph as lacking both written description and enablement. Applicants respectfully traverse.

Written description rejection of claims 1-14

The Office Action asserts that the broad genus of "a heterologous gene encoding pyruvate decarboxylase (pdc) or a functional equivalent thereof" is not adequately described in the specification. Amended claim 1, however, does not recite this broad genus. To advance prosecution, claim 1 has been amended to recite the genus of "a heterologous gene encoding pyruvate decarboxylase." The specification adequately describes this genus.

Heterologous genes encoding pdc were well known in the art prior to the filing of the application. A heterologous gene is a gene derived from a different species. (See Merriam Webster's definition of "heterologous"; Appendix A.) The specification teaches that heterologous pdc genes "may be from *Zymomonas sp*, preferably *Z. mobilis* or may be from yeast e.g. the *S. cerevisiae* pdc 5 gene." (Page 5, lines 1-2.) The nucleotide sequence of at least ten other pyruvate decarboxylase genes derived from species other than a species of gram-positive bacteria (*i.e.*, heterologous genes) were known in the art prior to the January 6, 2000

effective filing date of the application:

- Accession number X87929 (Appendix B) discloses the nucleotide and amino acid sequence for *Kluyveromyces lactis* pyruvate decarboxylase;
- Accession number X81854 (Appendix C) discloses the nucleotide and amino acid sequence for *Nicotiana tabacum* pyruvate decarboxylase;
- Accession number U71122 (Appendix D) discloses the nucleotide and amino acid sequence for *Arabidopsis thaliana* pyruvate decarboxylase-2;
- Accession number U65927 (Appendix E) discloses the nucleotide and amino acid sequence for *Neospora crassa* pyruvate decarboxylase;
- Accession number X92743 (Appendix F) discloses the nucleotide and amino acid sequence for *Oryza sativa* pyruvate decarboxylase;
- Accession number X59546 (Appendix G) discloses the nucleotide sequence and amino acid sequence for *Zea mays* pyruvate decarboxylase;
- Accession number U75311 (Appendix H) discloses the nucleotide and amino acid sequence for *Pichia stipitis* pyruvate decarboxylase;
- Accession number AF098293 (Appendix I) discloses the nucleotide sequence and amino acid sequence for *Aspergillus oryzae* pyruvate decarboxylase;
- Accession number AI728334 (Appendix J) discloses the nucleotide sequence for *Gossypium hirsutum* pyruvate decarboxylase isozyme 1; and
- Accession number Z54096 (Appendix K) discloses the nucleotide sequence of a cosmid from *Schizosaccharomyces pombe* chromosome 1 that includes the nucleotide and amino acid sequence of pyruvate decarboxylase (page 5).

These sequences were retrieved from the National Center for Biotechnology Information (NCBI) database and bear dates that are at least two and one half months before the priority date of the present application. A specification need not disclose what is well known to those skilled in the art and preferably omits that which is well known to those skilled and already available to the public. *In re Buchner*, 929 F.2d 660, 661 (Fed. Cir. 1991); *Hybritech, Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1384 (Fed. Cir. 1986). Thus, the teachings of the specification,

together with the knowledge in the art of numerous "heterologous genes encoding pyruvate decarboxylase," adequately demonstrates that applicants were in possession of the invention of claims 1-13 when the present specification was filed.

New claims 26-31 should also not be subject to this rejection. They each ultimately depend from claim 1 and thus also a recite gram-positive bacterium transformed with a heterologous gene encoding pdc.

Claim 14 recites particular deposited strains of bacteria. A reference in the specification to a deposit constitutes an adequate description of the deposited material sufficient to comply with the written description requirement. *Enzo Biochem v. Gen-Probe Incorporated*, 296 F.3d 1316, 1325 (Fed. Cir. 2002). Thus claim 14 also is adequately described.

Withdrawal of this rejection of claims 1-14 is respectfully requested.

<u>Enablement rejection of claims 1-14</u>

The Office Action asserts that claims 12-14 are not enabled because they recite novel vectors and strains of microorganisms, some of which have been deposited, but have not been indicated as being publicly available. (Paper 12, page 5, lines 10-16.) Applicants respectfully traverse.

Claim 12 recites a Gram-positive bacterium transformed with a plasmid, pFC1. The specification enables how to make and use plasmid pFC1. The specification discloses that plasmid pFC1 "was formed from a fusion of pAB124 (Bingham et al., Gen. Microbiol., 114, 401-408, 1979) and pUC18." (Page 12, lines 18-20.) The specification also teaches that as a result of fusing pAB124 and pUC18, plasmid pFC1 has a restriction enzyme cleavage map as shown in Figure 8. The plasmid pAB124 was known and used at the time the application was filed. Bingham (*J. Gen. Microbiol. 119*, 109-115, 1980) teaches, at Figure 2, a restriction

enzyme cleavage map of plasmid pAB124 (Page 114; Appendix L.) The sequence the plasmid pUC18 was known at the time the application was filed. GenBank accession number L09136 discloses the sequence of the complementary strand of pUC18. See Appendix M. One of skill in the art would have been able to make and use plasmid pFC1 as shown in Figure 8 using the well known structures of pAB124 and pUC18 and the routinely practiced techniques of molecular biology.

Claims 13 and 14 recite ten strains of bacteria. The recited bacterial strains are LN, LN-T (31), LN-T (32), TN, TN-P1, TN-P3, LN-S (J8), LN-D, LN-D11, and LN-DP1. Four of the recited bacterial strains, LN, TN, LN-S (J8), and LN-D, have been deposited with the National Collections of Industrial, Food and Marine Bacteria (NCIMB). These strains will also be made publicly available upon issuance of a patent. See the declaration of the inventors that accompanies this response. Appendix N. Thus each of these four bacterial strains are enabled.

The specification teaches one of skill in the art how to make and use the remaining six bacterial strains without undue experimentation. The specification teaches how to make and use bacterial strains TN-P1 and TN-P3. The specification discloses that bacterial strain TN-P1 is produced from bacterial strain TN after transformation with plasmid pBST22-zym. (Example 8; page 11, lines 21-24.) The specification discloses that bacterial strain TN-P3 is produced by transforming bacterial strain TN with plasmid pFC-PDC1. (Example 9, page 12, lines 15-18.) Those of skill in the art routinely practice transformation of bacteria. Thus, one of skill in the art could readily produce bacterial strains TN-P1 and TN-P3 using the deposited TN bacterial strain.

The specification also teaches one of skill in the art how to make and use bacterial strains LN-T (E31, E32). The specification teaches that bacterial strains LN-T (E31, E32) are "spontaneous transposon mutants from strain LN. Both strains are lactate deficient." (Page 9,

lines 6-7.) Thus strain LN-T (E31, E32) bacteria are strain LN bacteria that have spontaneously lost LDH activity. One of skill in the art would have been able to produce and identify LN-T strain bacteria by merely culturing the deposited LN strain bacteria and screening the cultured bacteria for LDH activity. Methods of assaying LDH activity of bacteria were also well known at the time the application was filed. Payton *et al.* (FEMS Microbiology Letters, 26, 333-336, 1985; Appendix O) teaches the identification of *Bacillus stearothermophilus* that lack LDH activity. Screening for *B. stearothermophilus* that lacked LDH activity "was achieved by selection for resistance to a suicide substrate, fluoropyruvate." (Page 333, lines 8-9 of the Summary; See also page 334, column 1, line 15 to column 2, line 40.)

The specification also teaches one of skill in the art how to make and use bacterial strain LN-D11 without undue experimentation. The specification discloses that LN-D11 bacteria are produced from strain LN-D bacteria "after repeated subculture." (Page 10, line 13.) LN-D11 bacteria are similar to strain LN-D bacteria except they are sensitive to kanamycin. See Figure 1, which discloses that LN-D strain bacteria are spo^-, ldh^-, km^R, (sporulation deficient, LDH recombination mutant, and kanamycin resistant) and LN-D11 bacteria are spo^- and ldh^- (sporulation deficient and LDH recombinant mutant). One of skill in the art, using the well known techniques of cell culture would have been able to subculture the deposited LN-D strain bacteria and screened it for kanamycin resistance without undue experimentation.

The specification also teaches one of skill in the art how to make and use bacterial strain LN-DP1. The specification discloses that LN-DP1 bacteria are "produced from strain LN-D11 after transformation with the replicative plasmid pBST22-zym." (Page 10, lines 24-25.) One of skill in the art could readily have produced bacterial strain LN-D11 as indicated above. One of skill in the art would further have been able to transform the LN-D11 strain bacteria with

plasmid pBST22-zym by routine experimentation. Thus, one of skill in the art would also have been able make and use bacterial strain LN-DP1 without having to resort to undue experimentation.

Withdrawal of this rejection to claims 12-14 is respectfully requested.

Rejection of claims 1, 2, 8, and 11 under 35 U.S.C. §102(b)

Claims 1, 2, 8, and 11 are rejected under 35 U.S.C. §102(b) as being anticipated by Danilevich *et al.* (*Molecular Biology*, 1994, 28(1):158-166; "Danilevich"). Applicants respectfully traverse.

To reject claims as anticipated, each and every element as set forth in the claim must be found either expressly or inherently described in a single prior art reference. *Verdegaal Bros. v. Union Oil Co. of California*, 814 F.2d 628, 631 (Fed. Cir. 1987). See also MPEP § 2131. Danilevich does not teach each and every element of claims 1, 2, 8, and 11.

Claim 1, the independent claim of the rejected set, is directed to a Gram-positive bacterium that has been transformed with a heterologous gene encoding pyruvate decarboxylase but that has native alcohol dehydrogenase function. Danilevich is cited as teaching "an identical Gram-positive bacterium belonging to the *Bacillus sp.* (Paper 12, page 6, lines 9-10.)

Danilevich teaches the construction of a recombinant plasmid that encodes *pdc*. The plasmid was transformed into a Gram-positive bacterium, *B. subtilis*. *B. subtilis*, however, lacks native alcohol dehydrogenase function. "The *B. subtilis* genome lacks both genes (*pdk* and *adh*)." (Page 105, second column, lines 14-15; See also page 109, first column, lines 29-30.) Thus Danilevich does not expressly or inherently teach a Gram-positive bacterium that has been transformed with a heterologous gene encoding pyruvate decarboxylase and that "has native

alcohol dehydrogenase function." Danilevich does not expressly or inherently teach each and every element of claim 1 or of dependent claims 2, 8, and 11 and does not anticipate these claims. Similarly, Danilevich does not anticipate new claims 26-31, which depend from claim 1.

Withdrawal of this rejection is respectfully requested.

Rejection of claims 1-8, 10, and 11 under 35 U.S.C. §103(a)

Claims 1-8, 10, and 11 are rejected under 35 U.S.C. §103(a) as being unpatentable over Danilevich in view of Hartley *et al.* (*Biotechnol.*, 1983, 145(1):390-396; "Hartley") and "the common knowledge in the art of molecular biology." (Paper 12, page 6, line 24.) Applicants respectfully traverse.

Each of claims 1-8, 10, and 11 recites a Gram-positive bacterium that (1) has been transformed with a heterologous gene encoding pyruvate decarboxylase and (2) has native adh function. To reject these claims as *prima facie* obvious the Patent Office must meet three criteria:

> First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations.

MPEP § 2143. The Patent Office has failed to meet the second criterion.

Danilevich is cited as teaching a Gram-positive bacterium with native *adh* activity that has been transformed with a plasmid that comprises a heterologous *pdc* gene. (Claim 1; Paper 12, page 6, lines 4-7.) Danilevich also is cited as teaching that the bacterium is a *Bacillus sp.* (claim 2), a heterologous *pdc* gene from *Z. mobilis* (claim 8), and transformation of the

bacterium with a plasmid comprising the heterologous gene (claim 11). (Paper 12, page 6, lines 9-11.) Hartley is cited as teaching the use of a thermophilic microorganism such as *B. stearothermophilus* for production of ethanol. (Claims 2-4; Paper 12, page 7, lines 14-15.) Hartley also is cited as teaching inactivation of the lactate dehydrogenase gene (claim 5) and a heterologous *pdc* gene from *Zymomonas sp.* or *S. cerevisiae* (claim 7). (Paper 12, page 8, lines 15-21.) The Office Action asserts:

> With the reference of Danilevich et al. which teaches the introduction of *Z. mobilis* pyruvate decarboxylase into *B. subtilis* and the reference of Hartley et al. which teaches the advantages of using a thermophilic *Bacillus* which has been genetically modified for alcohol production, it would have been obvious to one of ordinary skill in the art to transform *B. stearothermophilus* instead of *B. subtilis* as taught by Danilevich et al. and develop a strain for production of ethanol. With the common knowledge prevailing in the art of molecular biology, it would have been obvious to use methods in which the heterologous gene is introduced through a plasmid or incorporate it into the chromosome of the host bacterium. One of ordinary skill in the art would have been motivated to do so as Hartley et al. teach that use of thermophilic bacteria for production of ethanol has certain economic advantages over other fermentation methods. One of ordinary skill in the art would have a reasonable expectation of success since decarboxylase of *Z. mobilis* and the reference of Hartley et al. suggest the same and in addition provide methods to obtain lactate dehydrogenase non-producers.

Paper 12, page 8, lines 3-16.

As indicated above, Danilevich teaches that *B. subtilis* was transformed with a plasmid comprising the *pdc* gene. *B. subtilis* does not have native *adh* function. Hartley teaches manipulation of the genome of a thermophilic, Gram-positive bacterium that has a native *adh* gene (*B. stearothermophilus*) to inactivate the *lactate dehydrogenase* (*ldh*) gene. One of ordinary skill in the art, however, would not have combined these teachings to arrive at a Gram-

positive bacterium which has been transformed with a heterologous gene encoding pyruvate decarboxylase, and that has native alcohol dehydrogenase function.

Danilevich teaches that the metabolism of a bacterium having both pdc and adh gene function may be unpredictably altered: "Introduction of two actively expressed genes, *pdk* and *adh*, into *B. subtilis* may result in ethanol production or alter the cell metabolism in an unpredictable manner." (Page 109, second column, lines 33-36, emphasis added.) Thus, one of ordinary skill in the art would not have reasonably expected to be able to produce a bacterium comprising both an active, expressed heterologous pdc gene and native adh gene function.

Hartley does not remedy the deficiencies of Danilevich. Hartley teaches that "one might *try* a recombinant DNA approach to introduce and express the gene for pyruvate decarboxylase (PDC) from yeast or Zymomonas [*i.e.*, cells that have adh activity], which gives acetaldehyde + CO_2 from pyruvate (ADH then converts the acetaldehyde to ethanol)." (Page 896, line 41 to page 897, line 2, emphasis added.) But Hartley does not teach or suggest that a bacterium with both adh function and an active, expressed heterologous pdc gene could be successfully produced. In fact, if one of ordinary skill in the art looked to Danilevich for an expression vector encoding a pdc gene, she would have been taught that introduction of pdk and adh genes into a Gram-positive bacterium could alter its metabolism and would not reasonably have expected success.

In fact, in contrast to what would have been expected from the teachings of the cited art, the claimed bacteria have the unexpected properties of a significant increase in ethanol production and superior growth characteristics.

> Expression of *pdc* has resulted in a significant increase in ethanol production by the recombinant organism and has unexpectedly improved the organism's growth characteristics. Recombinant

> microorganisms, which prior to transformation with the *pdc* gene were highly unstable and difficult to culture, show significant increases in growth and survival rates both aerobically and anaerobically as well as an increase in the rate of ethanol production near to theoretical yields.

Page 4, lines 4-10.

One of ordinary skill in the art would have expected even less success at producing the bacterium of dependent claim 3. Claim 3 is directed to a <u>thermophilic</u> Gram-positive bacterium which is transformed with a heterologous gene encoding pyruvate decarboxylase that is expressed and active and which has native alcohol dehydrogenase function. Hartley teaches that a suitable vector for expression of pdc in a thermophilic bacterium was not yet available and, even if it <u>were</u> available a bacterium transfected with such a vector may not gain pdc function: "This strategy depends on developing suitable vectors for the thermophile and might fail if the enzyme were insufficiently thermostable." (Page 897, lines 2-4.)

Thus, according to Hartley, suitability of the vector is critical. Danilevich does not teach that the disclosed pdc expression vector is suitable for pdc expression in <u>thermophilic</u> bacteria. Danilevich teaches an expression vector that contains the pdc gene for use in Gram-positive bacterium. (Page 106, first column, lines 18-19.) But the vector was only tested in the Gram-positive bacterium *B. subtilis*, which is not a thermophile. Thus, the ordinary artisan would not have known whether use of the expression vector taught by Danilevich could satisfy either of the criteria set forth by Hartley: suitability for expression in thermophiles and stability of the expressed enzyme. Hartley's caveat regarding suitability of vectors, combined with Danilevich's teaching that a bacterium transfected with both adh and pdc genes may have altered metabolism and Danilevich's <u>lack</u> of teaching that the disclosed vector will express a functional pdc protein

in thermophilic bacteria, would not have provided the ordinary artisan with a reasonable expectation of success of producing the thermophilic bacterium recited in claim 3.

Withdrawal of this rejection to claims 1-8, 10, and 11 is respectfully requested.

Respectfully submitted,

Date: February 14, 2003

By: Michelle Holmes-Son

Michelle Holmes-Son
Registration No. 47, 660

BANNER & WITCOFF
1001 G Street, N.W., Eleventh Floor
Washington, D.C. 20001-4597
(202) 508-9100

Amendments to the Claims

C1

1. (Currently Amended) A Gram-positive bacterium which has been transformed with a heterologous gene encoding pyruvate decarboxylase, wherein the heterologous gene expresses an active pyruvate decarboxylase, ~~or a functional equivalent thereof, but~~ and wherein the bacterium has ~~solely~~ native alcohol dehydrogenase function.

2. (Original) A Gram-positive bacterium according to claim 1 wherein the bacterium is a Bacillus sp.

C2

3. (Currently Amended) A Gram-positive bacterium according to claim 1 wherein the bacterium is a thermophile.

4. (Currently Amended) A Gram-positive bacterium according to claim 2 wherein the Bacillus is selected from B. stearothermophilus; B. calvodax; B. caldotenax, B. thermoglucosidasius, B. coagulans, B. licheniformis, B. thermodenitrificans, and B. caldolyticus.

5. (Currently Amended) A Gram-positive bacterium according to claim 1 wherein the gene encoding lactate dehydrogenase expression has been inactivated.

6. (Original) A Gram-positive bacterium according to claim 5 in which the lactate dehydrogenase gene has been inactivated by homologous recombination.

7. (Previously Amended) A Gram-positive bacterium according to claim 1 in which the heterologous gene is from Zymomonas sp or from Saccharomyces cerevisiae.

8. (Original) A Gram-positive bacterium according to claim 7 in which the heterologous gene is from Z. mobilis.

9. (Currently Amended) A Gram-positive ~~bacteria according to claim 7~~ bacterium comprising a native adh gene and which has been transformed with a ~~in which the heterologous gene is~~ pdc gene from ~~S. cerevisiae~~ *S. cerevisiae*.

C3

10. (Previously Amended) A Gram-positive bacterium according to claim 9 wherein the heterologous gene is incorporated into the chromosome of the bacterium.

11. (Previously Amended) A Gram-positive bacterium according to claim 1 in which the bacterium has been transformed with a plasmid comprising the heterologous gene.

12. (Currently Amended) A Gram-positive bacterium ~~according to claim 11,~~ comprising a native adh gene and which has been transformed with a plasmid comprising a heterologous gene encoding pyruvate decarboxylase, wherein the plasmid is pFC1.

C4

13. (Currently Amended) A Gram-positive ~~bacteria according to claim 1~~ bacterium comprising a native adh gene and which has been transformed with a heterologous gene encoding pyruvate decarboxylase wherein the heterologous gene is operatively linked to the lactate dehydrogenase promoter from Bacillus strain LN (NCIMB accession number 41038).

C5

14. (Currently Amended) Strains LN (NCIMB accession number 41038); LN-T (E31, E32); TN NCIMB accession number 41039); TN-P1; TN-P3; LN-S (J8) (NCIMB accession number 41040); LN-D (NCIMB accession number 41041); LN-D11 and ~~LN-P1~~ LN-DP1.

C6

15-25. (Canceled)

26. (New) The gram-positive bacterium of claim 9 wherein the bacterium is a thermophile.

27. (New) The gram-positive bacterium of claim 12 wherein the bacterium is a thermophile.

28. (New) The gram-positive bacterium of claim 13 wherein the bacteria is a thermophile.

29. (New) The gram-positive bacterium of claim 9 further comprising an inactivated lactate dehydrogenase gene.

30. (New) The gram-positive bacterium of claim 12 further comprising inactivated lactate dehydrogenase gene.

31. (New) The gram-positive bacterium of claim 13 further comprising inactivated lactate dehydrogenase gene.

EXHIBIT A

Merriam-Webster

Merriam-Webster's COLLEGIATE® DICTIONARY

- Home
- Help
- Word of the Day
- Word Games
- Word for the Wise
- Books and CDs
- Company Info
- Customer Service
- Network Options
- Language Zone
- The Lighter Side
- Site Map

Merriam-Webster Unabridged
Member Log-in

Atlas | Reverse Dictionary | Rhyming Dictionary
Collegiate® Dictionary | Collegiate® Thesaurus | Unabridged

One entry found for **heterologous**.

Main Entry: **het·er·ol·o·gous**
Pronunciation: -'rä-l&-g&s
Function: *adjective*
Etymology: *heter-* + *-logous* (as in *homologous*)
Date: 1893
: derived from a different species <*heterologous* DNAs> <*heterologous* transplants>
- **het·er·ol·o·gous·ly** *adverb*

Get the Top 10 Most Popular Sites for "heterologous"
For More Information on "heterologous" go to Britannica.com

- **Search the Unabridged Dictionary on-line**
  and enjoy enhanced versions of Merriam-Webster's Collegiate® Dictionary and Thesaurus at Merriam-Webster Unabridged.
- **There's a word for you . . .**
  In fact, there's a new word waiting in your e-mail in box every morning when you subscribe to the Word of the Day.
- **Listen to Word for the Wise radio programs**
  Listen to broadcasts or read transcripts of previous Word for the Wise programs on public radio.

Pronunciation Symbols

\&\ as **a** and **u** in **abut**	\e\ as **e** in **bet**	\o\ as **aw** in **law**
\&\ as **e** in **kitten**	\E\ as **ea** in **easy**	\oi\ as **oy** in **boy**
\&r\ as **ur/er** in **further**	\g\ as **g** in **go**	\th\ as **th** in **thin**
\a\ as **a** in **ash**	\i\ as **i** in **hit**	\[th]\ as **th** in **the**
\A\ as **a** in **ace**	\I\ as **i** in **ice**	\ü\ as **oo** in **loot**
\ä\ as in **mop**	\j\ as **j** in **job**	\u\ as **oo** in **foot**
\au\ as **u** in **ut**	\[ng]\ as **ng** in **sing**	\y\ as **y** in **yet**
\ch\ as **ch** in **chin**	\O\ as **o** in **go**	\zh\ as **si** in **vision**

EXHIBIT B

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy

Search Nucleotide for Go Clear

Limits Preview/Index History Clipboard Details

Display default Save Text Add to Clipboard Get Subsequence

□ **1:** X87929. K.lactis PDC gene...[gi:3688421] Links

```
LOCUS       KLPDCGENE               3008 bp    DNA     linear   PLN 03-DEC-1998
DEFINITION  K.lactis PDC gene and upstream region of CYC1 gene.
ACCESSION   X87929
VERSION     X87929.1  GI:3688421
KEYWORDS    consensus sequence; CYC1 gene; PDC gene; promoter region; pyruvate
            decarboxylase.
SOURCE      Kluyveromyces lactis.
  ORGANISM  Kluyveromyces lactis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1  (bases 1 to 3008)
  AUTHORS   Ramil,E., Freire-Picos,M.A. and Cerdan,M.E.
  TITLE     Characterization of promoter regions involved in high expression of
            KlCYC1
  JOURNAL   Eur. J. Biochem. 256 (1), 67-74 (1998)
  MEDLINE   98417429
REFERENCE   2
  AUTHORS   Cerdan,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUN-1995) E. Cerdan, Univ. de la Coruna, Departamento
            de Biologia Celular y Mol., Facultad de Ciencias, Campus de La
            Zapateira s/n, E- 15071 La Coruna, SPAIN
REFERENCE   3  (bases 1 to 3008)
  AUTHORS   Cerdan,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-1995) M.E. Cerdan, Univ. de la Coruna,
            Departamento de Biologia Celular y Mol., Facultad de Ciencias,
            Campus de La Zapateira s/n, E- 15071 La Coruna, SPAIN
COMMENT     Related sequences L09727, X77316, X55905, X15668, X60834.
FEATURES             Location/Qualifiers
     source          1..3008
                     /organism="Kluyveromyces lactis"
                     /strain="NRRL-Y1140"
                     /db_xref="taxon:28985"
                     /chromosome="VI"
                     /clone="pEC1 and pART5"
     gene            426..2162
                     /gene="PDC"
     CDS             426..2162
                     /gene="PDC"
                     /codon_start=1
                     /product="pyruvate decarboxylase"
                     /protein_id="CAA61155.1"
                     /db_xref="GI:3688422"
                     /db_xref="SPTREMBL:O74722"
                     /translation="METKTLIHSGAAKEMSYTERYNVAPLIPLPEYLFHRLFQLNCRT
                     VFGVANYSTAKLYQAIAVSGIHGIQTINQLNTSFAVDAYGRAIGVSCYVTSESAELGH
```

```
                     VNGFFGSFCEYVPILQVVVLEQSHDLERLIGDVSIFHDVVDDPSEIDSCVRTLFWGKR
                     PVYMGLRSKDATKLVPSSSLNGNIADKMGIKNTFFQTDTIKRVIDKIIAEVYASSRPL
                     IVVDALIDRYNYNSTIQNFLTETGIPFVTTLMSKGSIDESLPNFVGTFLGTMSQPIVR
                     EYMNNADCTLILGCMIENFKNSYCRFNYKSKNQILLWNDRVKIENNIIPDILLHELLP
                     QLIASLDTTKIVNSRPVTIPNMIPRVEPQPVTFLRQEYLWFKMSTWLKQGDVIISESG
                     TSAIGLLQQKFPDNTRLVSQAIWNSSGYSIGACLGILAAYRDMGTLDKHRIILMVGDG
                     SLQFTFQELSTILTHGFKPYIFVINNQGYTVDRTLNREKTHLNATYFDIQPWELLKLP
                     SLFYSQEYFKRRCMSVGELNSLLSDKEFNKSDQLKIVELILPSMDVPVLLDPRDDSSD
                     DESSPQHKRPRT"
     gene            2517..3008
                     /gene="CYC1"
     misc_feature    2517..2543
                     /gene="CYC1"
                     /note="CRE1 consensus"
     misc_feature    2649..2660
                     /gene="CYC1"
                     /note="ROX1 consensus"
     misc_feature    2710..2721
                     /gene="CYC1"
                     /note="HAP 1 consensus"
     misc_feature    2799..2904
                     /gene="CYC1"
                     /note="HAP2/3/4/5 consensus"
     CDS             3006..>3008
                     /gene="CYC1"
                     /function="respiratory chain"
                     /codon_start=1
                     /product="cytochrome C"
                     /protein_id="CAA61156.1"
                     /db_xref="GI:4379104"
                     /translation="M"
BASE COUNT      963 a    588 c    549 g    908 t
ORIGIN
        1 aagttcattt caatgtcttt cggaaaacct atagttctat cctgtcccta atagtgttca
       61 tattgggcag tggattcttt ttatttatgt taagcattcc ctatgtacct cactaggaaa
      121 agggcacttg gacatcataa ttcttccttt tatagcaata cttttctttc tagcttccgt
      181 cctacctttt tcataccgga tttcaccttg tcagttcata ataatatcaa tactcaccta
      241 ttaagagcca catgacgtgc gtgcgattga gtacttaaaa actactgtta ccttttgtcc
      301 atcgtgtaac tttgcatttt ctttgtcaca ttttcctgta caagaactaa cttgacaaaa
      361 ttgattttcc ttgaagagaa gtcgaattaa aagtttaaca ggaatatact gggacggaca
      421 aaaagatgga aaccaagacc ctgatacact caggggctgc taaggagatg agctatactg
      481 agaggtataa tgtcgcaccg ctaatacctt tgcctgagta tctttttcat agattatttc
      541 agctcaactg ccgtactgtg tttggagttg ccaattattc aactgcgaaa ctatatcaag
      601 ccatagcagt tagcgggata catgggatcc aaacgataaa tcaattgaac acatcatttg
      661 cagtagatgc atacgggaga gctatcggag tgagctgcta cgtcactagc gaatctgcag
      721 aattaggcca tgttaatggt ttttttggat cattttgtga atatgttccc atcttacagg
      781 tagtcgtctt ggaacagtct catgatcttg agaggttgat tggagatgtt tcaatctttc
      841 atgacgtagt ggatgatcct tctgagatag acagttgtgt acgaacgctt ttttgggggga
      901 aacgccctgt ttatatgggc ctccgatcga aggatgccac gaaacttgtc cccagtagtt
      961 ctctaaacgg aaacatagca gataagatgg gcataaagaa cactttcttt caaacggaca
     1021 caataaagag ggttatagac aaaatcattg ccgaagttta cgcttcatca aggccattga
     1081 tcgtggttga tgcattaatt gatagatata attacaacag cacaatacaa aatttcctaa
     1141 cggaaacagg aataccattt gtgacaacct taatgtcaaa aggttctatt gacgaaagct
     1201 tacccaactt tgttgggact ttcttaggta ccatgtcaca gcccattgtt cgagaatata
     1261 tgaataatgc agattgtacg ttaattttag gatgtatgat cgaaaacttc aaaaactcgt
     1321 actgtagatt caactacaag agtaaaaacc aaatcttact ttggaatgat agagttaaaa
     1381 ttgaaaataa tataatacct gatattttac ttcatgaact actcccacag ctaatagcgt
     1441 cattagatac taccaaaata gtaaactctc gtccagtaac aataccgaat atgataccca
     1501 gagtagaacc gcaaccggtg acatttttgc gacaggagta cctttggttt aagatgtcta
     1561 catggctaaa acaaggtgac gttattattt ctgaatctgg tacctcagct attggtctcc
```

```
1621 tacaacaaaa attcccagat aatactaggc tagtatccca agcaatttgg aattcatctg
1681 gttattctat cggtgcatgc cttggaattt tggcagcata tcgcgatatg gggacattgg
1741 ataagcatag gattatatta atggtgggtg atgggtcttt acaatttact ttccaagaac
1801 taagcaccat attgacgcac ggtttcaaac cttacatttt tgtaataaat aatcaaggtt
1861 ataccgttga cagaaccttg aatagagaaa agacccatct gaacgctaca tacttcgata
1921 tccaaccgtg ggaactatta aaactcccgt cacttttcta ttcccaagag tatttcaaga
1981 gaagatgtat gtcggttggg gaattgaata gtttattaag cgataaagag ttcaataaat
2041 ccgatcagtt gaaaattgta gaactaatcc ttccatccat ggatgttcca gtactactcg
2101 atccacggga tgacagtagt gatgatgaat cttctccgca gcataaaaga cccagaacgt
2161 gaaaggagaa gggtgaaaca gatatcagag aactttcacg cttcaaatat atatattatt
2221 taataagcac tgaacacaca agtacctgct tggtgatcag tctcaatgtt aaccgcatta
2281 ccattaacaa cacggcacgc ttttcgtatc ggcactaacg aagagcacac caaaagtcac
2341 gtgtgcatta taccgatatt cacgcaaaga acagaatgaa gtaaataaat actctattga
2401 tgtcaaatac agataatatg agcgaaaaaa tatcatcttc tcagggtaac actgatcaga
2461 gatattccag aaccattacc attaagttaa ttgtcacgtg ccggtcagta cgtatcacgt
2521 gactgtaaca tacatcagcc aaccaatcag atcgcttccg ttcacagtga cgtccaacct
2581 ctgacaagaa ttgcttttcc ccacgcttta atccccatta agattaccct gcagcaaaac
2641 gcacaacggg aaaccaatga gaaaaaagtc tcagaggctt cacagaattt ggctcatctc
2701 gtgtcttgac cgatgttccc caaatgatag acttgtttcg aagcattttt tcggtttaag
2761 tagttaaaaa caggtaggag aaaaatttta gtatgcgcac tatggtttgt tgataatttc
2821 atatataact aacatattaa gatttgatat tcgattgaat ttgatgtttc tttagcaagt
2881 ctttttagcg ttttcttttt tctcctttct atcttaatta aagtttccca aactatcagt
2941 attaaggatt gattcgtcaa atcaaataac tgttacacaa ccaacacaac taatataaag
3001 tcataatg
//
```

Revised: July 5, 2002.

Oct 3 2002 17:48:23

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy

Search Nucleotide for Go Clear

Limits Preview/Index History Clipboard Details

Display default Save Text Add to Clipboard Get Subsequence

1: X81854. N.tabacum mRNA fo...[gi:551260] Links

```
LOCUS       NTMRNAPD1               1255 bp    mRNA    linear   PLN 06-NOV-1995
DEFINITION  N.tabacum mRNA for pyruvate decarboxylase 1.
ACCESSION   X81854
VERSION     X81854.1  GI:551260
KEYWORDS    pyruvate decarboxylase; tobPDC1.
SOURCE      common tobacco.
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1255)
  AUTHORS   Bucher,M., Brander,K.A., Sbicego,S., Mandel,T. and Kuhlemeier,C.
  TITLE     Aerobic fermentation in tobacco pollen
  JOURNAL   Plant Mol. Biol. 28 (4), 739-750 (1995)
  MEDLINE   95375236
REFERENCE   2  (bases 1 to 1255)
  AUTHORS   Mandel,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-SEP-1994) T. Mandel, Institute of Plant Physiology,
            Altenbergrain 21, 3013 Bern, FRG
FEATURES             Location/Qualifiers
     source          1..1255
                     /organism="Nicotiana tabacum"
                     /sub_species="samsun"
                     /db_xref="taxon:4097"
                     /tissue_type="leaves after anoxic treatment"
                     /dev_stage="adult"
     gene            1..1255
                     /gene="tobPDC1"
     CDS             <1..>1255
                     /gene="tobPDC1"
                     /EC_number="4.1.1.1"
                     /codon_start=3
                     /product="pyruvate decarboxylase"
                     /protein_id="CAA57447.1"
                     /db_xref="GI:551261"
                     /db_xref="SWISS-PROT:P51845"
                     /translation="AADGYARARGVGACVVTFTVGGLSVLNAIAGAYSENLPLICIVG
                     GPNSNDYGTNRILHHTIGLQDFSQEPRCFQTVTCYRAVVNNLEDAHELIDTAVSTALK
                     ESKPVYISIGCNLPGIPHPTFSREPVPFALSPRLSNMMGLEAAVEAAAEFLNKAVKPV
                     LVGGPKMRVAKASDAFVELSDACGYAVAVMPSAKGLFPEHHSHFIGTYWGAVSTAFCA
                     EIVESADAYLFAGPIFNDYSSVGYSLLLKKEKAIIVQPDRVTIGNGPAFGCVLMRDFL
                     AALAKRLKHNPTAFENYHRIYVPEGHPLKCEPKEALRVNVLFQHIQNMLSGDSVVIAE
                     TGDSWFNCQKLKLPKGCGYEFQMQYGSIGWSVGATLGYAQAAPEKRVIACIGDGSFQV
                     TAQDISTMLRCGQRTIIFLINNGGYT"
BASE COUNT      303 a    228 c    341 g    383 t
ORIGIN
```

```
   1 atgccgccga tgggtacgct agagcacgtg gtgtgggtgc gtgtgttgtt acgtttactg
  61 ttgggggatt aagtgttttg aatgctatag cgggtgctta tagtgagaat ttaccgttga
 121 tatgtattgt tgggggacct aattcgaatg attatggaac caacaggatt ttacatcata
 181 ctattgggtt gcaggatttt agtcaggagc ctcgttgctt tcaaactgtc acttgttacc
 241 gggctgtggt gaataattta gaggatgcac atgaactgat tgacactgct gtatctacag
 301 ctttgaaaga gagtaagccg gtgtatataa gcataggctg taatctgccg gggattccac
 361 accctacttt cagtcgtgaa ccagttccat ttgccctctc tcccagactg agtaatatga
 421 tgggtttgga agcagcagtg gaggcagcgg ctgagttctt gaacaaagca gtaaagccag
 481 tacttgttgg agggccaaaa atgcgggttg caaaggcatc tgatgctttt gttgagttgt
 541 ctgatgcctg tggttatgct gtcgcagtga tgccatctgc taaagggttg tttcccgaac
 601 accattcaca tttcattggg acttactggg gtgcagtgag cacagccttc tgtgctgaaa
 661 tcgtggaatc tgctgatgca tacttgtttg ctggacctat ttttaatgac tatagctctg
 721 ttggttattc tctgcttctc aagaaagaga aagcgatcat tgtccagcct gatcgtgtga
 781 ctattgggaa tggaccagcg tttggttgtg ttctgatgag ggatttcctt gctgcattag
 841 caaagagatt aaagcataat ccgactgctt ttgagaatta ccataggatt tatgttcctg
 901 agggacatcc tcttaaatgt gagcctaagg aggcattaag ggttaatgtt ctatttcaac
 961 acattcagaa tatgttgtct ggcgacagtg ttgtgattgc tgagacgggg gactcgtggt
1021 tcaattgcca gaagctgaaa ttgccaaaag gatgtgggta tgagttccaa atgcagtatg
1081 gatccatcgg ttggtctgtc ggtgcaactc ttggttacgc acaagctgca cctgaaaagc
1141 gggtgattgc ttgcatcggt gacggtagct ttcaggtgac tgctcaggat atttcaacga
1201 tgctgaggtg tggacaaaga accattatct tcttgataaa caatggtggc tatac
//
```

Revised: July 5, 2002.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

Oct 3 2002 17:48:23

EXHIBIT D

☐ **1:** U71122. Arabidopsis pyruv...[gi:1616786] Links

```
LOCUS       ATU71122                4069 bp    DNA     linear   PLN 08-OCT-1996
DEFINITION  Arabidopsis pyruvate decarboxylase-2 (Pdc2) gene, complete cds.
ACCESSION   U71122
VERSION     U71122.1  GI:1616786
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 4069)
  AUTHORS   Dolferus,R., Peacock,W.J. and Dennis,E.S.
  TITLE     Two pyruvate decarboxylase genes from Arabidopsis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4069)
  AUTHORS   Dolferus,R., Peacock,W.J. and Dennis,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-SEP-1996) Plant Industry, C.S.I.R.O., G.P.O. Box
            1600, Canberra, Act. 2601, Australia
FEATURES             Location/Qualifiers
     source          1..4069
                     /organism="Arabidopsis thaliana"
                     /strain="Landsberg-erecta"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /map="between DFR (7.5 cM) and RFLP m435 (1.8 cM)"
     gene            1447..3270
                     /gene="Pdc2"
     CDS             1447..3270
                     /gene="Pdc2"
                     /EC_number="4.1.1.1."
                     /codon_start=1
                     /product="pyruvate decarboxylase"
                     /protein_id="AAB16855.1"
                     /db_xref="GI:1616787"
                     /translation="MDTKIGSIDACNPTNHDIGGPPNGGVSTVQNTSPLHSTTVSPCD
                     ATLGRYLARRLVEIGVTDVFSVPGDFNLTLLDHLIAEPNLKLIGCCNELNAGYAADGY
                     ARSRGVGACVVTFTVGGLSVLNAIAGAYSENLPLICIVGGPNSNDYGTNRILHHTIGL
                     PDFTQELRCFQAVTCFQAVINNLEEAHELIDTAISTALKESKPVYISISCNLPAIPLP
                     TFSRHPVPFMLPMKVSNQIGLDAAVEAAAEFLNKAVKPVLVGGPKMRVAKAADAFVEL
                     ADASGYGLAVMPSAKGQVPEHHKHFIGTYWGAVSTAFCAEIVESADAYLFAGPIFNDY
                     SSVGYSLLLKKEKAIIVQPDRVTIGNGPAFGCVFMKDFLSELAKRIKHNNTSYENYHR
                     IYVPEGKPLRDNPNESLRVNVLFQHIQNMLSSESAVLAETGDSWFNCQKLKLPEGCGY
                     EFQMQYGSIGWSVGATLGYAQAMPNRRVIACIGDGSFQVTAQDVSPMIRCGHKTIIFL
                     INNGGYTIEVEIHDGPYNVIKNWNYTAFVEAIHNGEGKCWTPKVRCEEELVKAINTAT
                     NEEKESFCFIEVIVHKDDTSKELLEWGSRVSAANSRPPNPQ"
BASE COUNT     1105 a    828 c    855 g   1281 t
```

```
ORIGIN
        1 aagctttaaa ctttttccgt caatttcggt gtaaaacaca gagacataga agacagagag
       61 aaaaaacaca gacctgcgaa tttttcgtct ccctttcttc agctttcttc ttgtggcatt
      121 taacaacgag tcaataaagg gctaatgcga aaaaagacaa aaacatttga gagggccaaa
      181 aatattttaa ttacagcaaa aaaaattgtg gatttgtttg gttataacac tttctctgtt
      241 ttgttaaatt aatggtttgt ttttgtttgg gcttttagat aatatttttg cttaaagata
      301 tttccttctt gctttttat tatgcttctg tgatcagtta cttcttcttt tttagaggtc
      361 acgtgaaaac cacgtgtcta ctatgcaaag atctttgatt tatcaggcca ttcattagct
      421 aagcaatatt ataatcatcc ttcatctttt ggattagccc ggagataact actttaaaag
      481 cccattcaat catggactta aacttagtat tggtatatct acaaatttcg gttcaacttc
      541 attcagatta gttgctaaat tattgtttat ctgttctctg atacgttcgt tactttcata
      601 cattttcaac catacaataa accaaatcat atttcagatt tattccaaat atagtggacg
      661 acattcaatt tcccaaagta catttgattg aatgctcgag cttcttatct tatcccttga
      721 tccatcaaat tatcagcttc ttctaaccat tctatttcgt ttcacaatct tactttcttc
      781 ttttgcaagt aattaaagga agagcttgag gaattaggac aatcgaaata tttcttatca
      841 ttgcaaaagt caataacaaa accaaagctt ggtgaaagag attgaagaat ctgatgatcg
      901 agagagcgag aaagagattg tgggagatat gtgattgaag attggggatc tacggatttt
      961 tgtaatcgag tgagagcggc gggaaaatcg gcgttagggg ggaaatacga gtctctttag
     1021 ggtttcgaaa agtcgccaag agcgggagtt tctttttgag tcacgtggta ttgacacgaa
     1081 ccgggtcgga tcctatatcc aatcgggtac tatgttgccc ttgtgggcta cgcgtgtcgc
     1141 gtgagagcaa ctatctctgc ctactcacaa ctattctcaa cgtgaacttt ctaacttgtt
     1201 ctccaatatc caactatata aggacacgtg tctaatctca gagacacata aatatctgcg
     1261 gctaaaaata tcttattagc cgcctatttc ataccaccgt ttgattactt tagccaccgt
     1321 tgttttacgt ctataattaa gtatcttcca ctaatctcat attttcaaac tgtactttta
     1381 cctcgagctc tctccaattt tctcatagct ctattgttcg aaaattctga aacgaacgtt
     1441 accactatgg acactaagat cggatctatc gacgcgtgta acccgaccaa ccacgatatc
     1501 ggcggtcctc caaacggcgg agtctccacc gttcaaaaca caagtccact tcactccacc
     1561 accgtcagcc cctgcgacgc gactcttggc cgttacctag caagacggtt agtcgaaatc
     1621 ggcgtcaccg atgtcttctc cgttcctggt gatttcaacc tgacgcttct cgatcaccta
     1681 atcgccgaac caaacctcaa gctgatcggt tgctgcaacg agcttaacgc cggatacgct
     1741 gctgacggtt acgctagatc tcgcggtgtt ggtgcgtgcg tcgttacgtt caccgtcggt
     1801 ggattgagtg ttctgaatgc gatcgccggt gcttacagtg agaatctgcc cctgatttgc
     1861 atcgtcggtg gtccaaactc caacgattac ggtaccaata ggattcttca tcatacaatt
     1921 ggtttacctg atttcactca agagcttagg tgttttcaag ctgttacttg ttttcaagct
     1981 gtgattaata acttagaaga ggctcatgaa cttatcgata ctgcgatttc aactgctttg
     2041 aaagaaagca aacctgttta tatcagtatc agctgtaatt taccggcgat tcctcttccg
     2101 acgtttagtc gtcatcctgt tccgttcatg cttccgatga aggttagcaa tcagattggt
     2161 ttagatgcgg cggtggaggc agctgctgag ttcttgaaca aagctgtgaa gccagttctt
     2221 gttggtgggc cgaaaatgcg ggttgcgaaa gccgcggatg cttttgttga gcttgctgat
     2281 gcttctggct atggtcttgc tgtgatgcct tctgctaaag gacaagtacc tgagcatcac
     2341 aagcatttta tagggacgta ttggggagct gtgagtacag ctttttgtgc tgaaatcgtt
     2401 gaatctgcgg atgcttatct gtttgcaggt ccgattttca acgattacag ttctgttggg
     2461 tattctctgc ttctcaagaa ggagaaggca atcatcgttc agcctgatcg ggttactatc
     2521 ggtaacggac ctgcgtttgg atgtgttttt atgaaggatt ttctaagcga gttggctaaa
     2581 cgaattaagc acaacaacac ttcttatgag aattatcaca ggatctatgt cccagaagga
     2641 aagcctttga gagataaccc gaatgagtct ttgagggtta atgtactgtt ccaacacatt
     2701 cagaatatgc tctcttctga gtctgctgtg cttgctgaga caggagattc ctggttcaac
     2761 tgtcagaagc tgaagctccc tgaaggatgc ggttacgaat tccaaatgca gtacggatca
     2821 attggctggt cagtgggtgc tactctaggc tatgctcaag ccatgccaaa caggcgtgtc
     2881 attgcttgta ttggagatgg tagtttccag gtaaccgcac aggatgtatc tccgatgata
     2941 cggtgtgggc acaagaccat aatcttcctc atcaacaacg gaggctacac cattgaggtg
     3001 gaaattcacg atggtcctta caatgtcata aagaactgga actacacagc ttttgttgag
     3061 gccatacaca atggagaagg aaaatgctgg actcccaagg tgagatgcga ggaggagtta
     3121 gtgaaagcaa tcaacacggc aaccaatgag gaaaaagaga gcttttgttt cattgaagtg
     3181 atagtgcaca aagacgatac aagcaaggaa cttttggagt ggggctctag agtctctgct
     3241 gctaatagtc gtcccccaaa tccgcagtag agtatataaa tgcactcaac ttatatatat
     3301 ttcagatttg gtagtgtctt caccgttcta tgtaaagtag tgtgtatcct tttacaccca
     3361 tctggatgga aaaaaaatgt gtgtcccctt gaggatgtaa actgtttgag tgcctaaata
     3421 agaacgtgtt tcctctgctt catctttcgt cctcagtttt acgcagattt atcagagcta
     3481 tagtctttgt gtcttggatt ggtatgttga ttgacttaag aaagttgaga tccatcccaa
```

```
3541 gatgccaaat gattacttta aaggcagttg gtagagtcct ttcttttccc aatttatact
3601 aacttttgtt ttaccttttc ttgctttctg tttttttggt tcatggctaa atttgatctt
3661 gtttctatga ggaatgcagt gaggcgtctt gggagtgatc cactggttta caaaatactt
3721 catcgctatg ctgttgtaat gttacgaatc tttcgctgtt ggccataatt ttattactta
3781 agtaggatta aatagcttat tccatagtaa ctctcctaaa tttctgcata gagaaacctg
3841 tccaaatttg tgatgcagct tgaattcgac aggaattggg gttaatgcct caagtccggt
3901 gaagatgcag aaacccttgg actaactggc catgaactct acacaatcga cgtacctagt
3961 aatatcaatg agataaaacc aggccagtac ataactgtga ctactactga cactgctaaa
4021 tcttttgtct gcactttgcg gttagacaca tagttaatca caaaagctt
//
```

Revised: July 5, 2002.

Oct 3 2002 17:48:23

1: U65927. Neurospora crassa...[gi:1655908] Links

```
LOCUS       NCU65927                3741 bp    DNA     linear   PLN 01-NOV-1996
DEFINITION  Neurospora crassa pyruvate decarboxylase (cfp) gene, promoter
            region, and complete cds.
ACCESSION   U65927
VERSION     U65927.1  GI:1655908
KEYWORDS    .
SOURCE      Neurospora crassa.
  ORGANISM  Neurospora crassa
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariales; Sordariaceae; Neurospora.
REFERENCE   1  (bases 1 to 1699)
  AUTHORS   Temporini,E., Alvarez,M., Mautino,M., Kawaguchi,T., Kinghorn,J. and
            Rosa,A.
  TITLE     Cloning, sequence and functional analysis of the Neurospora crassa
            pyruvate decarboxylase gene promoter
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3741)
  AUTHORS   Temporini,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-1996) Esteban Temporini, Facultad de Ciencias
            Quimicas-Univ. Nac. Cordoba, Departamento de Quimica Biologica,
            Ciudad Universitaria s/n, Cordoba 5016, Argentina
FEATURES             Location/Qualifiers
     source          1..3741
                     /organism="Neurospora crassa"
                     /strain="wild type"
                     /db_xref="taxon:5141"
                     /chromosome="VIIR"
                     /clone="x10c4"
                     /tissue_type="mycelial"
                     /clone_lib="pMOcosX of Marc Orbach"
     protein_bind    456..460
                     /note="This is a sequence with high homology to the
                     binding site for GCR1 protein of Saccharomyces cerevisiae"
                     /bound_moiety="GCR1 protein"
                     /function="transcriptional activation of glycolytic genes"
     protein_bind    645..649
                     /bound_moiety="GCR1"
                     /function="transcriptional activation of glycolytic genes"
     protein_bind    741..753
                     /note="sequence with high homology to the binding site of
                     RAP1 protein of Saccharomyces cerevisiae"
                     /bound_moiety="RAP1"
                     /function="transcriptional regulation"
     protein_bind    1003..1015
                     /bound_moiety="RAP1"
                     /function="transcriptional regulation"
```

```
     protein_bind    1370..1382
                     /bound_moiety="RAP1"
                     /function="transcriptional regulation"
     protein_bind    1384..1396
                     /bound_moiety="RAP1"
                     /function="transcriptional regulation"
     protein_bind    1390..1394
                     /bound_moiety="GCR1"
                     /function="transcriptional activation of glycolytic genes"
     misc_signal     1602..1609
                     /function="transcription initiation signal in Neurospora
                     crassa"
     gene            1697..3409
                     /gene="cfp"
     CDS             1697..3409
                     /gene="cfp"
                     /codon_start=1
                     /evidence=experimental
                     /product="pyruvate decarboxylase"
                     /protein_id="AAB17969.1"
                     /db_xref="GI:1655909"
                     /translation="MVAQQQGKFTVGDYLAERLAQVGVRHHFVVPGDYNLILLDKLQA
                     HPDLKEVGCANELNCSLAAEGYARANGISACVVTYSVGALSAFNGTGSAYAENLPLVL
                     ISGSPNTNDPSQYHILHHTLGHPDYTYQYEMAKKITCCAVAIPRAIDAPRLIDRALRA
                     AILARKPCYIEIPTNLAGATCVRPGPISAITDPITSDKSALEAAAKCAAEYLDGKLKP
                     VILVGPKAGRAGSEKELIEFAEAMGCAVALQPAAKGMFPEDHKQFVGIFWGQVSSDAA
                     DAMVHWADAMICVGAVFNDYSTVGWTAVPNIPLMTVDMDHVTFPGAHFSRVRMCEFLS
                     HLATQVTFNDSTMIEYKRLKPDPPHVHTAEREEPLSRKEISRQVQEMLTDKTSLFVDT
                     GDSWFNGIQLKLPPGAKFEIEMQWGHIGWSIPAAFGYALRHPDRHTIVLVGDGSFQVT
                     AQEVSQMVRFKVPITIMLINNRGYTIEVEIHDGSYNKIKNWDYAMLVEAFNSTDGHAK
                     GLLANTAGELADAIKVAESHKEGPTLIECTIDQDDCSKELITWGHYVAAANARPPRNM
                     SVQE"
BASE COUNT      845 a   1139 c    943 g    814 t
ORIGIN
        1 ccgcggcctc ccactcagaa taacagtgcg cgtcatggta gaacaagcat atggggagca
       61 ggccttttgg caacaaaggt ggtcaatctc atggtcgtca ttctgtcaca tctcgcggtt
      121 gtgggaggag gttgagcaga cagaaaagta aatgggtaca cgaggggtgg tgggtgcctt
      181 ttgttggctg ggcgccaagg cttttccacc cctctctcca tagcagatat ggcacatgtg
      241 tttaactcaa taggctcatt tggtactata aatacaacct tacgccaaac ctttcatcga
      301 cgaaacttgt accaggaata tctatattct cattgacttt tcaaggacgg aggttccaat
      361 gatcagctct cgtcctaagc tcaaatgatc ttgtttggat tctatttggt aaagaagcag
      421 ccgccctcat ctccgttgcc ttggttgagg cggcacttcc gagggacaga ggtgaaagtg
      481 ttaggtcgca gtagggcgaa tcctcatttc caagtcagac aactttgaca tgatacgatg
      541 tgcaaactct tttcttctga atctctgctc ttatcttaag tttcttggaa tttcaaatac
      601 acaagagtgg ttgaatggag cttggtggtc tgttggttaa caggcttcca gcacttgggt
      661 tgcgtttctg actatcgccc tcactaaaaa catgttccat attttctatc ctcatgccct
      721 tggcactcag actcaagagc gcacccacca ccctcagggg acctaggcat agcttcagta
      781 tcgaactcat gtcaacgccg agcattgacc acatcgacct caaacctcaa caaacagaac
      841 aaccgacatg agctcagccc acgttgccat tcggcagacg gagccgaatc gtgctgactg
      901 ttttgaaaat ttaccgccgg atggtgccgt ctcggaaggc cagcctgagg catctgtgtc
      961 tggcatgtcg ccgacgcggt taaggtccgc ttttctgggg cgcaacccca ccagccgcca
     1021 atgggaaggg ccggtgtggc ttgggtgcga gtgcctggat tgcggcggtg gtctgatcaa
     1081 atcgtcgaat gctcgctgca tatcggaccg aacgaggccg cttatctttc cctacaacac
     1141 ttcatctcga ccgacgtcga acaaccggag cccagcagcg gatggccctg tgcaacgtcg
     1201 taatccatgg tcccccactc ctgtctatga tgcattgccc gcatgctctg acttcacaac
     1261 ctcgtaacag aatgttcaca ccaccagccg accgtggttg cgcttggtct gatggccctg
     1321 tcaagcaact ggcgaccgtc tggccattgc tcccaacgaa cgacaccatg cagccatgca
     1381 tcagcacacc ttccttgtcc cttggcccat gcggagacgc cggagcgtct tcggggacat
     1441 gtcatcgcgg agatggcctc aagccatcga tggcaatcct gtagcaggca ggccatcggg
     1501 acaagctctc tgagccaatc atcacatcac cgtggtcaac acaggacaat aaaaggcttc
```

```
1561 tcgcagcccg ccgtcctctt gggtaccgga cttctgcatg gtcatgaaag tacagcttcg
1621 acactaccac tactactacc acctccagtc tttgacactt cttcaccacc agatctcacc
1681 ctaaagagag agcatcatgg tagcccaaca acaaggaaag ttcacggtgg gcgactacct
1741 cgccgagcgt cttgctcagg tcggcgtccg ccaccacttt gtcgtccccg gcgactacaa
1801 cctcatcctc ctcgacaagc tacaagctca cccggatctc aaagaggttg gttgcgccaa
1861 cgagctcaac tgctctcttg ccgccgaagg ctatgcccgt gccaacggca tctccgcctg
1921 cgttgtcact tacagcgttg gtgccctctc ggccttcaac ggcaccggca gtgcctacgc
1981 agagaacctc cccttggtcc tcatcagcgg ctcgcccaac accaacgacc ccagtcaata
2041 tcacatcctt caccataccc tcggccaccc ggactacacc taccagtatg agatggccaa
2101 gaagattacc tgctgcgctg tggccattcc ccgcgccatc gatgcacccc gcctgatcga
2161 ccgtgccctc cgcgccgcca tccttgcccg gaagccctgc tacatcgaga tccccaccaa
2221 cctcgccgga gctacctgcg tgcgccccgg ccccatcagc gccatcaccg accccatcac
2281 cagcgacaag agcgctcttg aggccgccgc caagtgcgct gccgagtatc tcgacggcaa
2341 gctcaagccc gtcatcctgg ttggccccaa agctggccgt gccggttccg agaaagagct
2401 catcgagttc gccgaggcca tgggctgtgc tgttgccctg caacctgccg ccaagggcat
2461 gttccccgaa gaccacaagc agttcgtcgg catcttctgg ggacaagtta gctccgatgc
2521 cgcagacgcc atggtgcact gggccgacgc catgatttgc gtcggcgccg tctttaacga
2581 ctacagcacg gtcggatgga ctgccgtgcc caacatcccg ctcatgaccg tcgatatgga
2641 tcatgtcacc ttccccggcg cacacttcag ccgcgtccgg atgtgcgaat tcctcagcca
2701 cctggccacc caggtcacct tcaacgactc caccatgatc gagtacaagc gcctcaagcc
2761 cgaccccccg cacgtgcaca cggccgaacg cgaggagccg ctctcgcgca aggagatctc
2821 gcgacaggtg caggagatgt tgacagacaa gaccagcctg ttcgtcgaca cgggcgactc
2881 gtggttcaac ggcatccagc tgaagcttcc tcctggcgcc aagttcgaga tcgagatgca
2941 atggggccac atcggctggt cgatccctgc ggcctttggc tacgccttgc gccaccctga
3001 caggcacacc attgtgctgg tcggcgacgg atctttccag gtgacggccc aggaggtgtc
3061 gcagatggtg cgcttcaagg tacccatcac catcatgctc atcaacaacc gcggctacac
3121 gatcgaggtc gagatccacg acggttccta caacaagatc aagaactggg actacgccat
3181 gttggtggag gccttcaaca gcaccgacgg ccacgcaaag ggcctgctgg ccaacacggc
3241 cggcgagctc gctgacgcta tcaaggtggc tgagagccac aaggagggtc ctacgctgat
3301 cgagtgcacg attgatcagg atgactgcag taaggagctc atcacctggg gtcattacgt
3361 ggctgctgcg aatgcgaggc ctccccgcaa catgtcggtg caggagtaag gaagatgcta
3421 atgtggatag gcatgttgaa atgacggatg tgataccaaa tgatagtatt atcaaagata
3481 caacggcggc ggcggcggcc gaggtggtgg tagacaatgg gccacgatat agtgtatgga
3541 aaaaggtcaa acggggcttc aagtttttgt catgtcaata aatgttttgt tttcctttta
3601 ggctttctca tagatcgttc atctcatctc aatttctata tatctctctc tcctcattca
3661 tgtctctgtc ccacccggtt tggttaaaat ctatacaaag acaagtggta aacgaacaag
3721 aaaagctatt caatcgaatt c
//
```

Revised: July 5, 2002.

Oct 3 2002 17:48:23

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy

Search Nucleotide for Go Clear

Limits Preview/Index History Clipboard Details

Display default Save Text Add to Clipboard Get Subsequence

1: X92743. O.sativa mRNA for...[gi:1054820] Links

```
LOCUS       OSPDECOOH                856 bp    mRNA    linear   PLN 07-AUG-1997
DEFINITION  O.sativa mRNA for pyruvate decarboxylase.
ACCESSION   X92743
VERSION     X92743.1  GI:1054820
KEYWORDS    pyruvate decarboxylase.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Rivoal,J., Thind,S., Pradet,A. and Ricard,B.
  TITLE     Differential induction of pyruvate decarboxylase subunits and
            transcripts in anoxic rice seedlings
  JOURNAL   Plant Physiol. 114 (3), 1021-1029 (1997)
  MEDLINE   97377132
REFERENCE   2  (bases 1 to 856)
  AUTHORS   Ricard,B.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1995) B.C. Ricard, Institut National de la
            Recherche Agron., Station de Physiologie vegetale, B.P.81, F-33883
            Villenave d'Ornon Cedex, FRANCE
FEATURES             Location/Qualifiers
     source          1..856
                     /organism="Oryza sativa"
                     /cultivar="Cigalon"
                     /db_xref="taxon:4530"
                     /clone_lib="lambda ZAP"
                     /dev_stage="2-days-old seedlings subjected to 6 hours of
                     anoxia"
     CDS             <1..591
                     /EC_number="4.1.1.1"
                     /codon_start=1
                     /product="pyruvate decarboxylase"
                     /protein_id="CAA63404.1"
                     /db_xref="GI:1054821"
                     /db_xref="SPTREMBL:Q43005"
                     /translation="NILFKHIKELLSGDTALIAETGDSWFNCQKLRLPEGCGYEFQMQ
                     YGSIGWSVGATLGYAQAAKDKRVMSCIGDGSFQMTAQEVYTMLRCGQKSIIFLINNGG
                     YTIEVEIHDGPYNVIKNWDYTGLIDAIHNSDGNCWTKKVRTEEELIEAIATATGAKKD
                     CLCFIEIIVHKDDTSKELLEWGSRVSAANSRPPNPQ"
BASE COUNT      210 a    215 c    240 g    191 t
ORIGIN
        1 aacatcctct tcaagcacat caaggagttg ctgtccggcg acaccgcgct catcgccgag
       61 accggtgact cgtggttcaa ctgccagaag ctcaggctcc ccgagggctg cgggtacgag
      121 ttccagatgc agtacggctc gatcgggtgg tccgtcggcg ccacgctcgg ctacgcccag
      181 gcggccaagg acaagcgcgt catgtcctgc atcggtgacg gtagcttcca gatgacggcg
```

```
241 caggaagtgt acacgatgtt gcggtgcggg cagaagagca tcatcttcct catcaataat
301 ggtgggtaca ccatcgaggt ggagatccac gacgggccgt acaacgtcat caagaactgg
361 gactacaccg gcctcatcga cgcaatccac aactctgacg gcaactgctg gacaaagaag
421 gtccggactg aggaggagct gatagaggcg atcgcgacgg caacgggcgc caagaaagat
481 tgcctctgct tcatcgagat catcgtgcac aaagatgaca cgagcaaaga gcttctcgag
541 tggggatcca gggtatcggc tgccaacagc aggccgccga atccccagtg attttgctag
601 ctggtcttct catgacgcac actcgaatgc gctagtctgc caatgtccga gatatgctgc
661 ccattatcgt ctgtgaatta attgtctaat gcgtctgcat ccagttaaat tttctcagcc
721 gagaggttga agtatttatg tagatggggt tgcccgtctt atgattatgt ttgtaattca
781 atatgcttgg cattaatgtt ttgtaaacct gtaatgtatt gcgatgaaat aacatggacg
841 atttgaagga gaagaa
//
```

Revised: July 5, 2002.

Oct 3 2002 17:48:23

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy

Search Nucleotide for Go Clear

Limits Preview/Index History Clipboard Details

Display default Save Text Add to Clipboard Get Subsequence

1: X59546. Z.mays Pdc mRNA f...[gi:22394] Links

```
LOCUS       ZMPDCMRNA               4040 bp    DNA     linear   PLN 03-DEC-1998
DEFINITION  Z.mays Pdc mRNA for pyruvate decarboxylase.
ACCESSION   X59546
VERSION     X59546.1  GI:22394
KEYWORDS    anaerobic stress response; glycolytic enzyme; PDC gene; pyruvate
            decarboxylase.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3791)
  AUTHORS   Kelley,P.M., Godfrey,K., Lal,S.K. and Alleman,M.
  TITLE     Characterization of the maize pyruvate decarboxylase gene
  JOURNAL   Plant Mol. Biol. 17 (6), 1259-1261 (1991)
  MEDLINE   92032792
REFERENCE   2  (bases 1 to 3791)
  AUTHORS   Kelley,P.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAY-1991) P.M. Kelley, School of Biological Sciences,
            University of Nebraska-Lincoln, 245 Manter Hall, Lincoln, Nebraska,
            68588-0118, USA
FEATURES             Location/Qualifiers
     source          1..4040
                     /organism="Zea mays"
                     /strain="W22"
                     /db_xref="taxon:4577"
                     /clone="lambda clone 6"
                     /clone_lib="genomic EMBL3"
     TATA_signal     1261..1268
     prim_transcript 1294..3832
     gene            join(1415..1954,2045..2977,3084..3272,3367..3537)
                     /gene="Pdc gene"
     CDS             join(1415..1954,2045..2977,3084..3272,3367..3537)
                     /gene="Pdc gene"
                     /codon_start=1
                     /product="pyruvate decarboxylase"
                     /protein_id="CAA42120.1"
                     /db_xref="GI:22395"
                     /db_xref="SWISS-PROT:P28516"
                     /translation="METLLAGNPANGVAKPTCNGVGALPVANSHAIIATPAAAAATLA
                     PAGATLGRHLARRLVQIGASDVFAVPGDFNLTLLDYLIAEPGLTLVGCCNELNAGYAA
                     DGYARSRGVGACAVTFTVGGLSVLNAIAGAYSENLPVVCIVGGPNSNDYGTNRILHHT
                     IGLPDFSQELRCFQTITCYQAIINNLDDAHEQIDTAIATALRESKPVYISVSCNLAGL
                     SHPTFSRDPVPMFISPRLSNKANLEYAVEAAADFLNKAVKPVMVGGPKIRVAKAREAF
                     AAVADASGYPFAVMPAAKGLVPEHHPRFIGTYWGAVSTTFCAEIVESADAYLFAGPIF
                     NDYSSVGYSLLLKREKAVIVQPDRMVVGDGPAFGCILMPEFLRALAKRLRRNTTAYDN
```

```
                     YRRIFVPDREPPNGKPNEPLRVNVLFKHIKGMLSGDSAVVAETGDSWFNCQKLRLPEG
                     CGYEFQMQYGSIGWSVGATLGYAQAAKDKRVIACIGDGSFQVTAQDVSTMLRCGQKSI
                     IFLINNGGYTIEVEIHDGPYNVIKNWDYTGLVNAIHNSEGNCWTMKVRTEEQLKEAIA
                     TVTGAKKDCLCFIEVIVHKDDTSKELLEWGSRVSAANSRPPNPQ"
     exon            <1415..1954
                     /gene="Pdc gene"
                     /number=1
     intron          1955..2044
                     /gene="Pdc gene"
                     /number=1
     exon            2045..2977
                     /gene="Pdc gene"
                     /number=2
     intron          2978..3083
                     /gene="Pdc gene"
                     /number=2
     exon            3084..3272
                     /gene="Pdc gene"
                     /number=3
     intron          3273..3366
                     /gene="Pdc gene"
                     /number=3
     exon            3367..>3537
                     /gene="Pdc gene"
                     /number=4
     repeat_region   3579..3596
                     /note="tgcgtc repeat"
                     /rpt_type=TANDEM
     polyA_signal    3790..3795
     polyA_signal    3814..3819
BASE COUNT      898 a   1197 c   1003 g    942 t
ORIGIN
        1 ctcactgtcc atgccatcag atggtactgt acaggaagaa tgatactgta gttaagtacc
       61 cccaaaatat aaggcatttt caagattatt gaaaagtcaa agtgttgttt caagtttgac
      121 caaatttata taacgaagta ataacgttta cgatattaaa taaatgtata ccattagatt
      181 ctctattaaa tatattttta tactatatct actctgtgtc ataaatataa atatttatga
      241 ttctcgttat aattttaatc aaactttaaa aactttgact gtctaaattt tttgaaatga
      301 cttatatttt gggacagagg agagtaaggt tttagcctcg gcggcactgg aagaatgtta
      361 attagaagca tgcccacgtc atctttttgt tactgtattt atctaatgca atatcatgta
      421 ataaataaac acatatagct atatatattt ttgttaattt taagtgacga ccgaatagta
      481 gtttataaag agtaaagtac actaccgacc tacaaacttg gcatgcgggt aaacacgtac
      541 tcgaactttg tcagccaata aaaaaccatc caaaaccgag tttgtaaaac cagatgccga
      601 cgtggtcaga gccacttaag atgtctaact tggtttatat gttacaaagt tacatcttaa
      661 tttatcttta gcttccaaca tcctatctac atgtgtctag ctagagttac tcaccacact
      721 cgatcgggaa catggctaga gttgtcactg tatttctcta atgataatgt catgtaattt
      781 ttttctgtga ctgtcagtgg tcaccgaggt ttaagggagc aacttaattt aggacgggat
      841 ttttaaatac ggtataacat ttttggcacg cacaagcaca aaaccagtac cgccgcagtt
      901 tgttttggta caactgctga tactagcagc tacagtgaat cctatctgta gtgtgctggg
      961 gctatatgta gctgtagtgt gcaacggtaa tactaggaca tggaaacctg gaggcggcca
     1021 gctaggctag tactgtacta ctaggacaac agctcagctc aggcaacgcg ccggaaggca
     1081 cacgaagccc aagtttggtt tggtggtccg tttatcccag cccacagcat gcgaacagac
     1141 tggcccgcct tgccttggcg gcttggcact gtgtgcagtg taacaacctt tcgcccttcc
     1201 cctccctcag ggcttcgggc cttctgccct tctgccttca gcttcagacc cgttccgtgc
     1261 tatataaatc cccacgatgg gaggggttgc ccaccccacc catcatcagt cagtcagtca
     1321 gcagccgcaa cctcgtacaa tcacgaaagc agcacacgag tccacaacag caccccgcct
     1381 ctttttcttc acgtgcgttg cgtgccctgc ccccatggag accctcctgg ctggcaaccc
     1441 cgccaacggc gtggcgaagc cgacgtgcaa cggcgtcggc gccctgcccg tggccaactc
     1501 ccacgccatc atcgccacgc cggcggcggc tgcggcgacg ctggccccgg ctggcgccac
     1561 gctcggccgg cacctggcgc gccgcctcgt gcagatcggc gcctccgacg tcttcgccgt
     1621 cccgggggac ttcaacctca ccctgctcga ctacctcatc gccgagccgg gcctgaccct
```

```
1681 cgtcggctgc tgcaacgaac tcaacgccgg gtacgccgcc gacgggtacg cgcggtcgcg
1741 gggcgtcggc gcctgcgccg tcacgttcac cgtcggtggg ctcagcgtgc tcaacgccat
1801 cgccggcgcg tacagcgaga acctccccgt cgtctgcatc gtgggcgggc ccaactccaa
1861 cgactacggc accaaccgca tccttcacca caccatcggc ctccctgact tctcccagga
1921 gctccgctgc ttccagacga tcacctgcta ccaggtgctg cgcgctaata ctgctgtgtg
1981 cttctcatct tctctcttct gctagctgct cggcatcgct atatcttctc tctactctcg
2041 gcaggccatc atcaacaacc tggacgacgc gcatgagcag atcgacacgg ccatcgcgac
2101 ggcgctgagg gagagcaagc ccgtgtacat cagcgtcagc tgcaacctgg ccggcctctc
2161 ccacccgacc ttcagccggg atccggtgcc catgttcatc tcgccgaggc tgagcaacaa
2221 ggccaacctg gagtacgccg tggaggcggc ggccgacttc ctcaacaagg cggtgaagcc
2281 ggtgatggtg ggcgggccca agatccgggt ggccaaggcc agggaggcgt tcgccgccgt
2341 cgcggacgcc agcgggtacc ctttcgccgt catgccggcc gccaagggcc tggtgccgga
2401 gcaccacccg cgcttcatcg gcacctactg ggcgccgtc agcaccacct tctgcgccga
2461 gatcgtggag tccgccgacg cctacctctt cgccgggccc atcttcaacg actacagctc
2521 cgtcggctac tccctcctgc tcaagcggga gaaggccgtc atcgtgcagc ccgaccgcat
2581 ggtggtcggc gacggcccgg cgttcggctg catcctcatg cccgagttcc tccgcgcgct
2641 cgccaagcgc ctcaggcgca acaccacggc gtacgacaac taccgccgca tcttcgtccc
2701 cgaccgcgag ccgcccaacg gcaagcccaa cgagccgctt agggtcaacg tcctcttcaa
2761 gcacatcaag ggcatgctgt ccggcgactc cgcggtcgtc gccgagaccg gcgactcgtg
2821 gttcaactgc cagaagctca ggctccccga gggctgcggc tacgagttcc agatgcagta
2881 cggctccatc gggtggtcgg tcggcgcgac gctcgggtac gcccaggcgg ccaaggacaa
2941 gcgtgtcatc gcgtgcatcg gcgacggaag cttccaggta caatacggca catccatcca
3001 tggcacacat gctgtctctc tgtctgtgcc tgtacgtgct cactgctgat cggacgatcg
3061 acgacgtgtg gctggctgcc caggtgacgg cgcaggacgt gtccacgatg ctccggtgcg
3121 ggcagaagag catcatcttc ctcatcaaca acggcggcta caccatcgag gtggagatcc
3181 acgacggccc gtacaacgtg atcaagaact gggactacac gggcttggtc aacgccatcc
3241 ataactccga gggcaactgc tggaccatga aggtctgtcg atcagacaga ccctcctcca
3301 tccaacactg aaccccacca ccaccagctg cgcgctaatg aaaccgctcc accgtgtgcc
3361 ctgcaggttc ggacggaaga gcagctcaag gaagcgatcg ccacggttac gggcgccaag
3421 aaagactgcc tctgcttcat cgaggtaatc gtgcacaagg acgacacgag caaggagctc
3481 ctcgagtggg ggtccagggt ctccgccgcc aacagcaggc cgcccaaccc ccagtgatga
3541 gcccgtcggt cggtctcgag cgtttctgca atggcgcctg cgtctgcgtc tgcgtctgtc
3601 tgccctgaaa atttttcggg tcgtttactc gttttcagtc agttactcag tttgtccgtc
3661 ttcgtcgagc agaactacat ccctgctgag ggaagctacc aatctttatc ctacgatgat
3721 aatgtttgca atttcacttg cttccgattg agttctgtta taaatttttg caacattttt
3781 tctcttatca ataaaatttg tttttctctt atcaataaaa tttgttcgcg acgccggtct
3841 ccttagagac ctcggttact gtgcgccggg tgtgtatccc catgacctag acggacttaa
3901 gcctgaatcg tgtggaatag gtgatttagg gttagtttag caacccc ttt tttaagagat
3961 ttctattttt taagatagat tagtttattt tcctttgagg aaaacgaaat tccgtaagaa
4021 attatagtat taaactagcc
//
```

Revised: July 5, 2002.

Oct 3 2002 17:48:23

EXHIBIT H

1: U75311. Pichia stipitis p...[gi:2734882] Links

```
LOCUS       PSU75311                2305 bp    DNA     linear   PLN 19-FEB-1998
DEFINITION  Pichia stipitis pyruvate decarboxylase 2 (PsPDC2) gene, complete
            cds.
ACCESSION   U75311
VERSION     U75311.1  GI:2734882
KEYWORDS    .
SOURCE      Pichia stipitis.
  ORGANISM  Pichia stipitis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE   1  (bases 1 to 2305)
  AUTHORS   Lu,P., Davis,B.P. and Jeffries,T.W.
  TITLE     Cloning and characterization of two pyruvate decarboxylase genes
            from Pichia stipitis CBS 6054
  JOURNAL   Appl. Environ. Microbiol. 64 (1), 94-97 (1998)
  MEDLINE   98096791
   PUBMED   9435065
REFERENCE   2  (bases 1 to 2305)
  AUTHORS   Lu,P., Davis,B.P. and Jeffries,T.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-OCT-1996) Institute for Microbial and Biochemical
            Technology, USDA, FS, Forest Products Laboratory, One Gifford
            Pinchot Drive, Madison, WI 53705, USA
FEATURES             Location/Qualifiers
     source          1..2305
                     /organism="Pichia stipitis"
                     /strain="CBS 6054"
                     /db_xref="taxon:4924"
     TATA_signal     216..222
     TATA_signal     267..273
     gene            393..2102
                     /gene="PsPDC2"
     CDS             393..2102
                     /gene="PsPDC2"
                     /EC_number="4.1.1.1"
                     /function="decarboxylation of pyruvate during ethanol
                     fermentation"
                     /note="thiamine binding protein"
                     /codon_start=1
                     /product="pyruvate decarboxylase 2"
                     /protein_id="AAC03165.1"
                     /db_xref="GI:2734883"
                     /translation="MVSTYPESEVTLGRYLFERLHQLKVDTIFGLPGDFNLSLLDKVY
                     EVPDMRWAGNANELNAAYAADGYSRIKGLSCLVTTFGVGELSALNGVGGAYAEHVGLL
                     HVVGVPSISSQAKQLLLHHTLGNGDFTVFHRMSNSISQTTAFLSDISIAPGQIDRCIR
                     EAYVHQRPVYVGLPANMVDLKVPSSLLETPIDLKLKQNDPEAQEVVETVLKLVSQATN
                     PIILVDACALRHNCKEEVKQLVDATNFQVFTTPMGKSGISESHPRLGGVYVGTMSSPQ
```

```
                     VKKAVENADLILSVGSLLSDFNTGSFSYSYKTKNVVEFHSDYMKIRQATFPGVQMKEA
                     LQQLIKRVSSYINPSYIPTRVPKRKQPLKAPSEAPLTQEYLWSKVSGWFREGDIIVTE
                     TGTSAFGIIQSHFPSNTIGISQVLWGSIGFTVGATVGAAMAAQEIDPSRRVILFVGDG
                     SLQLTVQEISTLCKWDCNNTYLYVLNNDGYTIERLIHGKSASYNDIQPWNHLSLLRLF
                     NAKKYQNVRVSTAGELDSLFSDKKFASPDRIRMIEVMLSRLDAPANLVAQAKLSERVN
                     LEN"
     misc_feature    1734..1799
                     /gene="PsPDC2"
                     /note="encodes thiamine binding region"
BASE COUNT      681 a    472 c    460 g    692 t
ORIGIN
        1 aagaattgct tcttaccttg atggcattcc taatgaaacc atccactcaa taggctgtgt
       61 tgaatgctgc ctaatcgcat atcctacaca taataggaat gtctccagct tatggacgat
      121 taaggaacca ttcatgcgtt taaacatgct gcacattaat cttttctaca agttcatcac
      181 gaactgatgg gaaatgcctt cagatgccca gggcctataa tcgacaactg tactagcgaa
      241 taagcagctg atcacgattt acttcttata aataagtgca aatcgtcgcg caccatatca
      301 acttttggca gccattgcac aacaattttt cacttagaat atctacacct atttttaact
      361 gtagatccca ataatatact gtccttatag caatggtatc aacctaccca gaatcagagg
      421 ttactctagg aaggtacctc tttgagcgac tccaccaatt gaaagtggac accattttcg
      481 gcttgccggg tgacttcaac ctttccttat tggacaaagt gtatgaagtt ccggatatga
      541 ggtgggctgg aaatgccaac gaattgaatg ctgcctatgc tgccgatggt tactccagaa
      601 taaagggatt gtcttgcttg gtcacaactt ttggtgttgg tgaattgtct gctttaaacg
      661 gagttggtgg tgcctatgct gaacacgtag gacttctaca tgtcgttgga gttccatcca
      721 tatcgtcaca ggctaaacag ttgttgctcc accatacctt gggtaatggt gacttcactg
      781 tttttcacag aatgtccaat agcatttctc aaactacagc atttctctca gatatctcta
      841 ttgcaccagg tcaaatagat agatgcatca gagaagcata tgttcatcag agaccagttt
      901 atgttggttt accggcaaat atggttgatc tcaaggttcc ttctagtctc ttagaaactc
      961 caattgattt gaaattgaaa caaaatgatc ctgaagctca agaagttgtt gaaacagtcc
     1021 tgaagttggt gtcccaagct acaaaccccca ttatcttggt agacgcttgt gccctcagac
     1081 acaattgcaa agaggaagtc aaacaattgg ttgatgccac taattttcaa gtctttacaa
     1141 ctccaatggg taaatctggt atctccgaat ctcatccaag attgggcggt gtctatgtcg
     1201 ggacaatgtc gagtcctcaa gtcaaaaaag ccgttgaaaa tgccgatctt atactatctg
     1261 ttggttcgtt gttatcggac ttcaatacag gttcattttc atactcctac aagacgaaga
     1321 atgttgttga attccactct gactatatga aaatcagaca ggccaccttc ccaggagttc
     1381 aaatgaaaga agccttgcaa cagttgataa aaagggtctc ttcttacatc aatccaagct
     1441 acattcctac tcgagttcct aaaaggaaac agccattgaa agctccatca gaagctcctt
     1501 tgacccaaga atatttgtgg tctaaagtat ccggctggtt tagagagggt gatattatcg
     1561 taaccgaaac tggtacatct gctttcggaa ttattcaatc ccattttccc agcaacacta
     1621 tcggtatatc ccaagtcttg tggggctcaa ttggtttcac agtaggtgca acagttggtg
     1681 ctgccatggc agcccaggaa atcgacccta gcaggagagt aattttgttc gtcggtgatg
     1741 gttcattgca gttgacggtt caggaaatct ctacgttgtg taaatgggat tgtaacaata
     1801 cttatcttta cgtgttgaac aatgatggtt acactataga aaggttgatc cacggcaaaa
     1861 gtgccagcta caacgatata cagccttgga accatttatc cttgcttcgc ttattcaatg
     1921 ctaagaaata ccaaaatgtc agagtatcga ctgctggaga attggactct ttgttctctg
     1981 ataagaaatt tgcttctcca gataggataa gaatgattga ggtgatgtta tcgagattgg
     2041 atgcaccagc aaatcttgtt gctcaagcaa agttgtctga acgggtaaac cttgaaaatt
     2101 gaattaaaca taccaattca ccgctattta gttattatat tcatagcagc agttatagaa
     2161 ttggcttatt atgtgttcgt agtaattgag tgagctgttt catttattgt acttatagct
     2221 agaaagagaa ccaggacgtt actaaaacga ctaaacttgt tcaaagtaaa atatattgaa
     2281 aattccattc agaatgaatt ttacc
//
```

Revised: July 5, 2002.

Oct 3 2002 17:48:23

1: AF098293. Aspergillus oryza...[gi:4323052] Links

```
LOCUS       AF098293                2525 bp    DNA     linear   PLN 06-APR-1999
DEFINITION  Aspergillus oryzae pyruvate decarboxylase (pdcA) gene, complete
            cds.
ACCESSION   AF098293
VERSION     AF098293.1  GI:4323052
KEYWORDS    .
SOURCE      Aspergillus oryzae.
  ORGANISM  Aspergillus oryzae
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE   1  (bases 1 to 2525)
  AUTHORS   Lee,D.W., Koh,J.S., Kim,J.H. and Chae,K.-S.
  TITLE     Cloning and nucleotide sequence of one of the most highly expressed
            genes, a pdcA homologue of Aspergillus nidulans, in Aspergillus
            oryzae
  JOURNAL   Biotechnol. Lett. 21, 139-142 (1999)
REFERENCE   2  (bases 1 to 2525)
  AUTHORS   Lee,D.W., Koh,J.S., Kim,J.H. and Chae,K.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-1998) Faculty of Biological Sciences, Chonbuk
            National University, Dukjin-dong 1-ga, 664-14, Chonju, Chonbuk
            561-756, Republic of Korea
FEATURES             Location/Qualifiers
     source          1..2525
                     /organism="Aspergillus oryzae"
                     /db_xref="taxon:5062"
     gene            <282..>2182
                     /gene="pdcA"
     mRNA            join(<282..369,404..484,566..1790,1846..2053,
                     2153..>2182)
                     /gene="pdcA"
                     /product="pyruvate decarboxylase"
     CDS             join(282..369,404..484,566..1790,1846..2053,2153..2182)
                     /gene="pdcA"
                     /codon_start=1
                     /product="pyruvate decarboxylase"
                     /protein_id="AAD16178.1"
                     /db_xref="GI:4323053"
                     /translation="MSLSTSSGDFVRLAFVQYTVFLVSINAPAWDYNLVALDYLPKCD
                     LHWVGNCNELNAGYAADGYARINGMSALVTTFGVGELSALNAIAGAYSEFVPIVHIVG
                     QPHTKSQKDGMLLHHTLGNGDFNVFTRMSADISCTLGCLNSTHEVATLIDNAIRECWI
                     RSRPVYISLPTDMVTKKIEGERLDTPLDLSLPPNDPEKEDYVVDVVLKYLHAAKKPVI
                     LVDACAIRHRVLDEVHEFVEKSGLPTFVAPMGKGAVDETHKNYGGVYAGTGSNPGVRE
                     QVESSDLILSIGAIKSDFNTTGFSYRIGQLNTIDFHSTYVRVRYSEYPDINMKGVLQK
                     IVQRMGNLNVGPVSPPSNLLPDNEKASTEQAITHAWLWPTVGQWLKEKDVVITETGTA
                     NFGIWDTRFPAGVTAISQVLWGSIGYSVGACQGAALAAKEQGRRTVLFVGDGSFQLTL
                     QEVSTMIRNNLNPIIFVICNEGYTIERYIHGWEAVYNDIQPWDFLNIPVAFGAKDKYK
```

```
                     GYKVTTRDELRELFANEEFASAPCLQFEIDSRIGR"
BASE COUNT       659 a    600 c    587 g    679 t
ORIGIN
        1 aagcttgtaa ttttggcgta ctatctgtac tatctaccca tgacatcacg ttatcatgac
       61 gtatggcctg gtcttactgt tgcacgtcga gttgtacagg acattctgta ttcactacta
      121 tataactgcc ccccttgtc catgtttcca ggaagaattc agtatccaac aatccataac
      181 aacaacctta tcatctacta taccttccat tttccattcc tcttactgag tccttacaat
      241 ggcgacagat atcgctacaa gggatcttcg caagcccata gatgtcgctg agtacctctt
      301 caggcgactt cgtgaggttg gcgttcgtgc agtacacggt gttcctggta agcatcaatg
      361 cgccagcttg taccgcaact actgacagta tcaattcgca taggggacta caacttggtg
      421 gctttggatt acctgccaaa atgcgatctt cattgggtag gaaactgtaa tgagcttaat
      481 gccggtatgt ctccactgaa tggtcaatgg gaattctggg tatggtaaaa gagaggcgca
      541 tatcgctaac ggtggtcaaa tgcaggatac gctgctgatg gatacgctcg aatcaatgga
      601 atgtctgctt tagtcaccac ctttggtgtg ggtgagctat cggcgctcaa tgctattgct
      661 ggtgcatact ccgaatttgt gcctatcgtt cacattgttg gtcaaccgca tacgaaatca
      721 cagaaagatg gaatgctcct ccaccacacc ttgggcaacg gcgacttcaa cgtcttcacc
      781 agaatgagtg ccgacatctc ttgcacactt ggatgtttga actcaactca cgaagtggcg
      841 accctcattg ataatgctat ccgagaatgt tggattcgta gtcgaccggt ttatatctct
      901 ctccctaccg atatggtgac aaagaaaatc gagggagaac ggctggatac ccctctcgat
      961 cttagtctac caccgaacga tcccgaaaaa gaagattacg ttgtggatgt ggttctcaag
     1021 tatctgcacg ctgcaaagaa acccgttatt cttgtcgatg cttgtgctat ccgccatcgt
     1081 gtgctcgatg aagttcatga gttcgtggaa aaatctgggc tacccacatt cgtggctcca
     1141 atgggtaaag gagcagtgga tgagactcac aagaactacg gcggtgttta cgctggtact
     1201 ggatcaaacc caggtgttcg tgagcaagtc gaatcttcag acttgattct gagcatcggt
     1261 gctatcaagt ccgatttcaa cacgactggg ttctcttacc gtattggcca actcaacacc
     1321 attgacttcc atagtacata cgtgcgcgtc cggtactccg aataccctga tatcaacatg
     1381 aaaggcgtcc ttcaaaagat tgttcaaaga atgggcaatc tcaatgtcgg accagtctcg
     1441 ccgccgtcga acctactgcc ggacaacgag aaggcatcaa ccgaacaggc gattacccac
     1501 gcatggctct ggcctactgt cgggcagtgg ctgaaagaaa aggatgttgt tatcacggaa
     1561 accggcactg ccaatttcgg tatctgggac actcggttcc cggcaggtgt tacagccatt
     1621 agtcaggttc tttggggtag tatcggctat tcagttggag cttgtcaagg tgctgcgttg
     1681 gccgcaaaag agcagggccg acggactgta cttttcgtgg gtgacggaag tttccagctg
     1741 acgctccagg aagtcagcac catgataaga aataacctta accctatcat gtaagtatcg
     1801 cctttgggtg acacatgtct catgcccgac taacatcgta cgtagctttg tcatttgcaa
     1861 cgaaggatat accatcgaac ggtacattca tggatgggaa gctgtttaca atgacatcca
     1921 gccctgggac ttcttgaaca ttcctgtggc attcggcgcg aaggacaagt acaaaggata
     1981 caaggtcaca acccgagacg agttgaggga gcttttcgca aatgaagaat ttgcttcggc
     2041 accctgtctc caggtaggtt tcatttcgtt taccccctgtc ggaaaagtat cagtggtgct
     2101 aaccaaagca gttggttgag ctccacatgc ctcgcgacga ttgcccagcc agtttgaaat
     2161 tgacagccga atcggccgct gagcggaaca agtcccttta atttcttcaa gacatcctcc
     2221 aggatagacc agatctcttt ccttctttac ggctcgggaa atcggtttca gttagcatgc
     2281 caccatgcac acggggcatt tttaattcca tcatgacatg tggttaagag attttgttcg
     2341 tttgcgatag ttttataatg agttctttat ctgctttccc tctccttctt tcctttcaat
     2401 accttctaag ttaaaacaat aatcagattt aattaaataa caatgaatta aatgttaatt
     2461 tagagagaga gacagagaaa aagagaaaaa gaaataacca gactctctcg ggtgggaggg
     2521 tgaac
//
```

Revised: July 5, 2002.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

Oct 3 2002 17:48:23

1: AI728334. BNLGHi10479 Six-d...[gi:5047186]

IDENTIFIERS

```
dbEST Id:       2637583
EST name:       BNLGHi10479
GenBank Acc:    AI728334
GenBank gi:     5047186
```

CLONE INFO

```
Clone Id:       (5')
DNA type:       cDNA
```

PRIMERS

```
Sequencing:     T3 Primer
PolyA Tail:     Unknown
```

SEQUENCE

```
GAGAAAGCAATCATAGTACAACCAAATCGTGTGACGATAGGCAATGGTCCTTCTTTTGGC
TGGGTTTTCATGGCTGACTTCTTAAGTGCATTGGCAAAGAAACTTAAGAAAAATACCACT
GCCGTGGAGAACTATCGGCGAATCTTTGTCCCTCCAGGCATGCCTTTGAAGCATGGGAAT
GACGAGCCTCTGAGAGTCAATGTTCTTTTTAAGCACATTCAGGATATGCTAAGCAAAGAC
TCTGCGGTTATTGCTGAAACTGGAGACTCATGGTTCAATTGTCAGAAGCTCCATCTTCCA
GAGAACTGCGGTTATGAATTCCAGATGCAGTATGGATCCATTGGTTGGTCAGTTGGTGCC
ACGCTTGGATATGCTCAGGCAGCTAAAGGCAAACGTGTGATTGCTTGCATTGGCGATGGT
AGTTTCCANGTGACAGCTCAAGAATATTCCAACAATGATCCCATGTGGACAAAAAAGCAT
CATATTTCTTATCAATAATGGAGGCTATACAATTGAAGTTGAGATTCATGATGGCCCTTA
CAATGTTATCAAGAACTGGAATTACACTGCCTTTGTTGATGCCATCCACAATGGTGAAGG
CAAATGCTGGACAGCCAAGGTGAGGACCGANGATGAACTGATAAATGCT
```

```
Entry Created:  Jun 11 1999
Last Updated:   Jun 11 1999

PUTATIVE ID     Assigned by submitter
                PYRUVATE DECARBOXYLASE ISOZYME 1 (PDC)
```

LIBRARY

```
Lib Name:       Six-day Cotton fiber
Organism:       Gossypium hirsutum
Cultivar:       Acala Maxxa
Tissue type:    immature fiber
Develop. stage: Six days post anthesis
Lab host:       XL1-Blue
Vector:         pBluescript II KS+
```

SUBMITTER

```
Name:           Ben Burr
Lab:            Biology Department
Institution:    Brookhaven National Laboratory
```

Address: Upton, NY 11973, USA
Tel: 516-344-3396
Fax: 516-344-3407
E-mail: burr@bnlux1.bnl.gov

CITATIONS

Title: ESTs from developing cotton fiber
Authors: Blewitt,M., Matz,E.C., Davy,D.F., Burr,B.
Year: 1999
Status: Unpublished

MAP DATA

Revised: July 5, 2002.

Oct 3 2002 17:48:23

EXHIBIT K

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy

Search Nucleotide for Go Clear

Limits Preview/Index History Clipboard Details

Display default Save Text Add to Clipboard Get Subsequence

1: Z54096. S.pombe chromosom...[gi:984221] Links

```
LOCUS       SPAC13A11              12168 bp    DNA     linear   PLN 18-OCT-1999
DEFINITION  S.pombe chromosome I cosmid c13A11.
ACCESSION   Z54096
VERSION     Z54096.1  GI:984221
KEYWORDS    cyp51; cytochrome P450; cytosol aminopeptidase; Fes/CIP4 homology
            domain; pyruvate decarboxylase; Rho GTPase protein; Thiamine
            pyrophosphate enzymes; ubiquitin carboxy-terminal hydrolase.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 12168)
  AUTHORS   Hunt,S., Devlin,K., Churcher,C.M., Barrell,B.G., Rajandream,M.A.
            and Walsh,S.V.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-SEP-1995) Schizosaccharomyces pombe chromosome I
            sequencing project, Sanger Centre, Hinxton Hall, Hinxton, Cambridge
            CB10 1RQ E-mail: barrell@sanger.ac.uk
COMMENT     Notes:
            Details of yeast sequencing at the Sanger Centre are available on
            the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/S_pombe)
            Protein coding regions (CDS) have been predicted with the help of
            computer analysis using the Genefinder program in PomBase (an ACEDB
            database) with additional predictions for the branch-acceptor sites
            supplied by the program Sp3splice. CAUTION: It is possible that for
            any individual CDS we may have underestimated or overestimated the
            number of introns/exons or we may not have chosen the correct
            splice donor/acceptor sites.
            CDS are numbered using the following system eg SPAC5H10.01c. SP (S.
            pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c
            (complementary strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous.
            The length in codons is given for each CDS.
            IMPORTANT: This sequence MAY NOT be the entire insert of the
            sequenced clone.  It may be shorter because we only sequence
            overlapping sections once, or longer, because we arrange for a
            small overlap between neighbouring submissions.
            Cosmid c13A11 overlapped at the 5' end by comsmid c2F7 and at the
            3' end c3H8.
FEATURES             Location/Qualifiers
     source          1..12168
                     /organism="Schizosaccharomyces pombe"
                     /strain="972h-"
                     /db_xref="taxon:4896"
```

```
                     /chromosome="I"
                     /map="IL"
                     /clone="cosmid c13A11"
     gene            complement(join(1..340,385..487))
                     /gene="SPAC13A11.01c"
                     /note="SPAC2F7.18c"
     CDS.            complement(join(1..340,385..487))
                     /partial
                     /gene="SPAC13A11.01c"
                     /note="SPAC13A11.01c, len:146, SIMILARITY:Saccharomyces
                     cerevisiae, YFE7_YEAST, hypothetical 82.2 kd protein in
                     EMP47-SEC53 intergenic region, (714 aa), fasta scores:
                     opt: 263 E():2.3e-09, (29.5% identity in 146 aa overlap)"
                     /codon_start=1
                     /product="putative Rho GTPase protein"
                     /protein_id="CAA90802.1"
                     /db_xref="GI:984222"
                     /db_xref="SWISS-PROT:Q09697"
                     /translation="MISSFSNGFWSKDYATGVKKLFDCLDNGVEENEQVKNLLKLYKE
                     ANEEFGEKLQEITKECLKGKKPENTEDGATSNKAFEGLRSEIANQGKQHIRIAKDLET
                     LIIAPFSKMSIDHSQKLQTSQQVLTNQIKSYEKKYYTLKKTKSAY"
     misc_feature    1..104
                     /note="overlap with cosmid c2F7, positions 42407 42510
                     EM:Z50142 S.pombe chromosome I cosmid c2F7"
     misc_feature    complement(join(144..340,385..469))
                     /gene="SPAC13A11.01c"
                     /note="Match to PF00611 FCH, Fes/CIP4 homology domain
                     Score 103.18"
     misc_feature    complement(341..354)
                     /gene="SPAC13A11.01c"
                     /note="ctaacaacgaatag, splice branch and acceptor"
     misc_feature    complement(379..384)
                     /gene="SPAC13A11.01c"
                     /note="gtattt, splice donor sequence"
     gene            complement(1251..2738)
                     /gene="SPAC13A11.02c"
                     /note="cyp51"
     CDS             complement(1251..2738)
                     /gene="SPAC13A11.02c"
                     /note="SPAC13A11.02c, len:495"
                     /codon_start=1
                     /product="probable cytochrome p450 51"
                     /protein_id="CAA90803.1"
                     /db_xref="GI:984223"
                     /db_xref="SWISS-PROT:Q09736"
                     /translation="MAFSLVSILLSIALAWYVGYIINQLTSRNSKRPPIVFHWIPFVG
                     SAVAYGMDPYVFFRECRAKYGDVFTFVCMGRKMTAFLGVQGNDFLFNGKLADLNAEEA
                     YSHLTTPVFGKDVVYDIPNHVFMEHKKFIKSGLGFSQFRSYVPLILNEMDAFLSTSPD
                     FGPGKEGVADLLKTMPVMTIYTASRTLQGAEVRKGFDAGFADLYHDLDQGFSPVNFVF
                     PWLPLPRNRRRDRAHKIMQKTYLKIIKDRRSSTENPGTDMIWTLMSCKYRDGRPLKEH
                     EIAGMMIALLMAGQHTSAATIVWVLALLGSKPEIIEMLWEEQKRVVGENLELKFDQYK
                     DMPLLNYVIQETLRLHPPIHSHMRKVKRDLPVPGSKIVIPANNYLLAAPGLTATEEEY
                     FTHATDFDPKRWNDRVNEDENAEQIDYGYGLVTKGAASPYLPFGAGRHRCIGEQFAYM
                     HLSTIISKFVHDYTWTLIGKVPNVDYSSMVALPLGPVKIAWKRRN"
     misc_feature    complement(1263..2642)
                     /gene="SPAC13A11.02c"
                     /note="Match to PF00067 p450, Cytochrome P450 Score
                     169.72"
     misc_feature    complement(1407..1436)
                     /gene="SPAC13A11.02c"
```

```
                     /note="PS00086 Cytochrome P450 cysteine heme-iron ligand
                     si gnature"
     gene            join(4637..4702,4797..5363)
                     /gene="SPAC13A11.03"
     CDS             join(4637..4702,4797..5363)
                     /gene="SPAC13A11.03"
                     /note="SPAC13A11.03, len:210, SIMILARITY:Arabidopsis
                     thaliana, CAB43931, hypothetical 14.7 kd protein., (128
                     aa), fasta scores: opt: 236, E():2.7e-08, (37.5% identity
                     in 120 aa)"
                     /codon_start=1
                     /label=SPAC13A11.03
                     /product="hypothetical coiled-coil protein"
                     /protein_id="CAA90804.1"
                     /db_xref="GI:984224"
                     /db_xref="SWISS-PROT:Q09739"
                     /translation="MPPKGLSLAEKRRRLEAIFHDSKDFFQLKEVEKLGSKKQIVLQT
                     VKDVLQSLVDDNIVKTEKIGTSNYYWSFPSDAKRSRESVLGSLQAQLDDLKQKSKTLD
                     ENISFEKSKRDNEGTENDANQYTLELLHAKESELKLLKTQLSNLNHCNPETFELKNEN
                     TKKYMEAANLWTDQIHTLIAFCRDMGADTNQIREYCSIPEDLDDLQLPIL"
     misc_feature    4703..4708
                     /gene="SPAC13A11.03"
                     /note="gtaagt, splice donor sequence"
     misc_feature    4775..4796
                     /gene="SPAC13A11.03"
                     /note="ctaatctatgtttttttatag, splice branch and acceptor"
     gene            complement(join(5836..5905,5951..6015,6069..6366,
                     6412..7099,7200..7332,7411..7506))
                     /gene="SPAC13A11.04c"
     CDS             complement(join(5836..5905,5951..6015,6069..6366,
                     6412..7099,7200..7332,7411..7506))
                     /gene="SPAC13A11.04c"
                     /note="SPAC13A11.04c, len:449,
                     SIMILARITY:Schizosaccharomyces pombe, O74442, ubiquitin
                     carboxyl-terminal hydrolase, (457 aa), fasta scores: opt:
                     536, E():3.1e-27, (30.6% identity in 320 aa)"
                     /codon_start=1
                     /label=SPAC13A11.04c
                     /product="ubiquitin carboxyl-terminal hydrolase"
                     /protein_id="CAA90805.1"
                     /db_xref="GI:984225"
                     /db_xref="SWISS-PROT:Q09738"
                     /translation="MPGDVEGCQHLKLKPADVENYQKICTQIFSCHFVPRRCSTCKRI
                     NKRSIRCLSCHSVGCLWGHHGEEHAMEHTHMIGVDVKNGHTYCFGCQDYVYQTELETL
                     RFKIKNIKAWQSDHKRLPEKYNQMVCLEAYRKYPPVCATAGLRGIQNLGATCFMSVIL
                     QSILHNPLVRNLFFSGFHTSTDCKRPTCMTCAIDDMFSSIYNSKNKSTFYGPTAVLNL
                     MWKLSKSLCGYSQQDGHEFFVYLLDQMHTESGGGTSMPCTCPIHRIFSGSLKNVVTCL
                     DCKKERVAVDPLMDISLDINEPTLQGCLERFVSKEKVQYSCHSCGSKNAIKQLVFDKL
                     PPTICMQLKRFEQNNFAMSTKIDKQVSYPAFLRMRYNFNQDDVDYQLYSVVCHKGTLD
                     TGHYIAYTYYQNQWFLLDDTTIVEVKESEVLNSQAYLLFYHERQILYSDEMTVKTEN"
     misc_feature    complement(5906..5920)
                     /gene="SPAC13A11.04c"
                     /note="ctaaccagattatag, splice branch and acceptor"
     misc_feature    complement(5945..5950)
                     /gene="SPAC13A11.04c"
                     /note="gtaggc, splice donor sequence"
     misc_feature    complement(6016..6028)
                     /gene="SPAC13A11.04c"
                     /note="ctaattttgtag, splice branch and acceptor"
     misc_feature    complement(6063..6068)
```

```
                     /gene="SPAC13A11.04c"
                     /note="gtatat, splice donor sequence"
     misc_feature    complement(6069..6149)
                     /gene="SPAC13A11.04c"
                     /note="Match to PF00443 UCH-2, Ubiquitin carboxyl-terminal
                     hydrolases family 2 Score 53.18"
     misc_feature    complement(6367..6385)
                     /gene="SPAC13A11.04c"
                     /note="ttaacaaatattattttag, splice branch and acceptor"
     misc_feature    complement(6406..6411)
                     /gene="SPAC13A11.04c"
                     /note="gtaaaa, splice donor sequence"
     misc_feature    complement(7018..7071)
                     /gene="SPAC13A11.04c"
                     /note="Pfam match to entry PF00442 UCH-1, Ubiquitin
                     carboxyl-terminal hydrolases family 2"
     misc_feature    complement(7100..7110)
                     /gene="SPAC13A11.04c"
                     /note="ctaattcttag, splice branch and acceptor"
     misc_feature    complement(7194..7199)
                     /gene="SPAC13A11.04c"
                     /note="gtacgt, splice donor sequence"
     misc_feature    complement(7261..7278)
                     /gene="SPAC13A11.04c"
                     /note="PS00190 Cytochrome c family heme-binding site
                     signat ure"
     misc_feature    complement(7333..7352)
                     /gene="SPAC13A11.04c"
                     /note="ttaacatttctttcatttag, splice branch and acceptor"
     misc_feature    complement(7405..7410)
                     /gene="SPAC13A11.04c"
                     /note="gtaagg, splice donor sequence"
     gene            8608..10149
                     /gene="SPAC13A11.05"
     CDS             8608..10149
                     /gene="SPAC13A11.05"
                     /note="SPAC13A11.05, len:513, SIMILARITY:Bos taurus,
                     AMPL_BOVIN, cytosol aminopeptidase, (487 aa), fasta
                     scores: opt: 1008, E():0, (42.1% identity in 442 aa)"
                     /codon_start=1
                     /label=SPAC13A11.05
                     /product="cytosol aminopeptidase"
                     /protein_id="CAA90806.1"
                     /db_xref="GI:984226"
                     /db_xref="SWISS-PROT:Q09735"
                     /translation="MKGLGLSTRTFNWSSLSSILLPRIPLATTKADSLILAVRHDKQV
                     FSEDYRQVVDQYFETSPKKNDIRLFWNTQGFVRLAIVQLEENVSEKSVRSAAAEAAKI
                     LKSNGAKSIAVDGMGFPKDAALGAALATYDFSLRRDHLSVYQDEKVVEKENLFTSPAP
                     ERLTFQLLSNTSEKKTATAEENAFKVGLIEAAAQNLARSLMECPANYMTSLQFCHFAQ
                     ELFQNSSKVKVFVHDEKWIDEQKMNGLLTVNAGSDIPPRFLEVQYIGKEKSKDDGWLG
                     LVGKGVTFDSGGISIKPSQNMKEMRADMGGAAVMLSSIYALEQLSIPVNAVFVTPLTE
                     NLPSGSAAKPGDVIFMRNGLSVEIDNTDAEGRLILADAVHYVSSQYKTKAVIEASTLT
                     GAMLVALGNVFTGAFVQGEELWKNLETASHDAGDLFWRMPFHEAYLKQLTSSSNADLC
                     NVSRAGGGCCTAAAFIKCFLAQKDLSFAHLDIAGVMDKQLNSWDCDGMSGRPVRTIIE
                     VARKY"
     misc_feature    8743..10047
                     /gene="SPAC13A11.05"
                     /note="Match to PF00883 Peptidase_M17, Cytosol
                     aminopeptidase family Score 764.41"
     misc_feature    9685..9708
```

```
                     /gene="SPAC13A11.05"
                     /note="PS00631 Cytosol aminopeptidase signature"
     gene            11713..12168
                     /gene="SPAC13A11.06"
                     /note="SPAC3H8.01"
     CDS             11713..12168
                     /partial
                     /gene="SPAC13A11.06"
                     /note="SPAC13A11.06, len:151,
                     SIMILARITY:Schizosaccharomyces pombe, O42873, putative
                     pyruvate decarboxylase, (570 aa), fasta scores: opt: 709,
                     E():0, (71.2% identity in 153 aa overlap)"
                     /codon_start=1
                     /product="pyruvate decarboxylase"
                     /protein_id="CAA90807.1"
                     /db_xref="GI:984227"
                     /db_xref="SWISS-PROT:Q09737"
                     /translation="MSGDILVGEYLFKRLEQLGVKSILGVPGDFNLALLDLIEKVGDE
                     KFRWVGNTNELNGAYAADGYARVNGLSAIVTTFGVGELSAINGVAGSYAEHVPVVHIV
                     GMPSTKVQDTGALLHHTLGDGDFRTFMDMFKKVSAYSIMIDNGNDAAEKI"
     misc_feature    11755..12162
                     /gene="SPAC13A11.06"
                     /note="Match to PF00205 TPP_enzymes, Thiamine
                     pyrophosphate enzymes Score 168.62"
     misc_feature    12046..12168
                     /gene="SPAC13A11.06"
                     /note="615 (+1) 1 123 SPAC3H8 EM:Z69086 S. pombe
                     chromosome 1"
BASE COUNT     3926 a   2206 c   2256 g   3780 t
ORIGIN
        1 tagtaagcgc tttttgtttt tttaagtgta tagtatttct tttcataact ctttatttga
       61 ttggtaagaa cttgttggct ggtttgtaac ttctgagaat gatcaataga cattttagag
      121 aaaggagcaa taattagcgt ttctaagtcc ttggcaatac gaatatgttg ctttccttga
      181 ttagcaatct cggagcgcaa cccttcaaaa gccttgttag aagtagcacc gtcttctgta
      241 ttctctggct ttttcccttt tagacattct ttcgttattt cttgaagctt ttcaccaaat
      301 tcctcattcg cttctttata taacttcaac aagtttttta ctattcgttg ttagaggaag
      361 tcctatatta atttatttaa ataccttgct cattctcttc aactccatta tccaggcaat
      421 cgaaaagttt tttcacacca gtcgcgtagt ctttactcca aaaaccattt gaaaagctcg
      481 atatcatttc aattcgagac tgaaactgtc acaacctgtc aagaaatata gaggtggttg
      541 acatgtctac gccctttatt actcaactta agtaaatagc gaattgctac gcagaaaagt
      601 tttacccatc aaggatgcca aaagtatcca gtaaataaga caacagatac gttacttgac
      661 acacataaga tagcatacca ctaatccttt ataaaatgta aacttcattc agttgaatag
      721 cagcctacgc ttttacccac tttgtattat taatgctatg gttggatcga tggtacaatt
      781 aaaaactaca aaattttaat taatagctta ctgagaaacc ttgcgtttgc acgatgaggt
      841 tttgaaaaga gagagtagcg tactttattg ttataatgcc tacatgaaaa cgaaagagag
      901 cgtcccgacg ttattaaaac tataatatcg taaaaatgta tcactaatga tgacagggta
      961 gttttcatat cgatttagtg cgaggataag gaagtttgaa attgcataat acattctttt
     1021 gacatttgaa atgctaccaa atagtcattt aaaacgctaa aatatttaca atttcataaa
     1081 acaataaagc acttaaaatt taaaatgtct agcaaatttc atggttattc gttactcgtc
     1141 agataaaaca aaaacgttta acacgggcat gaaaagagaa aggcggaacc aaagttagaa
     1201 accaatacca taagcgtatt gcaaatgcat tcaaattagg ggaaattccc ttaattacga
     1261 cgcttccaag caattttaac aggacccaag ggaagagcaa ccatactgga atagtcaacg
     1321 ttgggaactt ttccaataag agtccaggtg taatcgtgta caaacttgct aataatagtg
     1381 ctcaaatgca tgtaggcaaa ttgctcacca atgcaacgat gacgaccagc tccaaagggt
     1441 aggtatggag aagcagcacc ctttgtaacg agaccataac cataatcaat ttgctcagca
     1501 ttttcgtctt cattgacgcg gtcattccag cgcttggggt caaagtcagt agcatgagtg
     1561 aagtattcct cttcagtagc agttaaaccg ggagcagcaa gaagatagtt gttggcaggg
     1621 ataacaatct ttgaaccagg gacgggtaag tcacgtttga ctttacgcat atgactgtgg
     1681 atagggggat ggagtctaag agtttcttga ataacatagt tgagaagagg catgtcctta
     1741 tattggtcaa acttcaattc caagttttca cctacaacac gcttctgctc ttcccacaac
```

```
1801 atttcaatga tttcaggttt gctacccaaa agagcaagaa cccaaacgat agtggcagca
1861 gaggtgtgtt gaccagccat caaaagtgca atcatcatac cagcaatttc atgttccttc
1921 aaaggacggc catctctata cttgcagctc atcaaagtcc aaatcatatc agtaccagga
1981 ttttcagtgc tagagcgacg atctttaata attttaagat aagtttttg cataattttg
2041 tgggctctat cacggcgtct gtttctaggg agaggtaacc aagggaaaac aaagttaaca
2101 gggctaaatc cttgatcaag gtcatggtac aagtcagcga aaccagcgtc aaagcccttg
2161 cgaacctcag caccttgtaa agtacgagaa gcagtgtaaa tagtcatgac aggcatggtc
2221 ttcagcaaat cagcaacgcc ttccttacca ggtccgaagt caggagaagt gctgagaaaa
2281 gcatccattt cgttaagaat gagggggaca taagagcgga attgagagaa gccaagacca
2341 gatttgataa acttcttgtg ctccataaaa acatggttgg gaatatcgta aacgacatct
2401 ttaccgaata caggggtagt taaatgagag taagcttctt cagcattcaa atcagctaat
2461 ttaccgttaa agaggaaatc attaccttgg acaccaagga aagcggtcat tttacgaccc
2521 atgcaaacaa aggtgaaaac atcaccatat ttggcacgac actcgcgaaa aaatacataa
2581 gggtccatac cgtatgcgac agcagaacct acaaagggta tccaatggaa gacaattggg
2641 gggcgtttgc tattgcgaga ggtaagttgg ttaataatat aaccaacata ccaagccaat
2701 gcaatcgata ataaaatcga aactaaagag aatgccatcc tgacttatga aaaaaaaagt
2761 aatgctcaag tgggtaaagc aggtaaaaaa taatcagacg aaaaaaataa aatcagtcag
2821 taaagtagta aggggtaagc agatagcaac caaaaacgaa taagaccaac caaaaaacaa
2881 cgacaacaac aacaaatgtg taacagataa aaatagttca accctttgag ctaatgcaaa
2941 taaacttgaa aaagaaaagg agaaagaaaa aatgaattga ttaacacgca gctacgagaa
3001 caacgaatga aacaaataga aaaaagcgtt aaaaacgaaa tcgaaaacag acaataatat
3061 gaaagaaaaa caaaaaaaaa ttacgtttcc tttcacttaa cgaatgagat tgggattttg
3121 aaaagacaat tttcaaaaac aacgatcgat acagaaattt tagatcttgt ctttttaag
3181 gtacttctac aagtgaagaa tcgatttatc ttctctttga caccaaatac gctggtttcg
3241 acgacgattc aagctaactt gggtctcgcg ctgactgatg ggagttggta gaattatata
3301 tgcaacataa tggccgtgta accgtttgtt tatgctgaaa atccgtaaaa aaacatgcta
3361 cttgatcact tcaaaattac actttttat ccaatcatat ttttcagata ctaccttatt
3421 gttctttctg gtcaaggtga attagccact aagccaatga gaaatccaat tttgttacct
3481 gagaataggg cgacatcgta cgattcgagg gtttgtgagg tgaatgactg cggggcttat
3541 gttgcgatcg aacgatggca cgtccattgg accgagataa tacacatgtt taccgtagct
3601 cgtaacgcac atttaacaaa ctacttgctt ctaccctact ttatcaccta aaaataaaaa
3661 tgtgtcaaca tgtgcacgtc tactatacct tctactattt gttcggtaaa ccaccaaccg
3721 tgcctcctta tacgtcttct ccaaacatgt aattggattc gaacaacttt agctcgacaa
3781 gtcgtttagc ttacgaacat ccaaaagctt atctaatgac ctaactaact cgctcttttt
3841 ggtcgcatga ctttgcatac gattgtttag taataaaaac aataatcctt ttcacacctg
3901 tttttaccca tgggaaagcg gttcccaaag tttgtcactt actcccatat tctaatctag
3961 attaaaacac atttagctgc gaacatggca caccggttaa tacccagcca tagttgaaat
4021 tagtagacaa attgactgag tatctcttat caaaagtcaa tacacatttt gtttgcgctg
4081 acattaaaca tcgttgctct tatgaatgct atgtgtgtgt gtgtgtgtat gtggaaatgt
4141 gttagtcagt cagtgtttgc ttgaaattcg ctgtgtgaaa ataataaagt gaaaaataaa
4201 taaaaggttg taaaaccgaa ttcgtacagt aaggtgagcg ccgacaaaag ctcttgagaa
4261 aaaaaaaatc atgctgctat cagtattcga gtcttttcga tttcgtggca aaaggaagtt
4321 ttcctcacat gtttttgtgc gacaatagcc aaccccattc acttaattcc ttaatttcct
4381 caagggtcaa aaagattagg atagctgaat aacgttttgc attgcattat gttgcaatac
4441 gatagtaaga cattataggg tatatatcat agttggaaag tgcggatgtg cattttagca
4501 gactacgtta tcattcagta aatgtttggg atctaaactc tcatttctcc aaaagattag
4561 aaagcatcta cttcttgcat cacaacaaac ttttaaactc catatttgtt caaactaatt
4621 aataaatttt taaacaatgc ctcccaaggg actatcgctt gcagagaaaa gaagaagatt
4681 agaggccatt ttccatgact ctgtaagttc gtgatatgat tgttcaagtg ccaaaacatg
4741 atttttgccc attcgtgtct cactttccac attgctaatc tatgtttttt ttatagaaag
4801 attttttca actaaaggag gtggaaaagc tgggctctaa aaagcaaatc gttttgcaaa
4861 ctgtaaagga cgttttacag tctctcgttg acgacaatat cgtcaaaacg gaaaagattg
4921 gcactagcaa ttactactgg tcttttccct ccgatgctaa acgttcaaga gagtcagttc
4981 tcggatctct tcaagctcaa cttgatgatt tgaaacaaaa atctaaaact ttagatgaaa
5041 acattagttt cgagaaatcc aaaagggata acgaaggaac cgaaaatgat gccaatcaat
5101 acactttgga actgctacat gcaaaagaat cagaactaaa acttcttaaa acacaacttt
5161 caaatttaaa tcactgtaat cctgaaacct ttgaattaaa aaatgagaat actaaaaagt
5221 acatggaagc tgccaatttg tggaccgatc aaatccatac tttgatagct tttgtcgag
5281 acatgggtgc cgataccaat caaattcgtg agtattgcag cattccagag gatttggacg
5341 atctgcaact accgattttg tgaacttgac atgcttcctt aattcaaaga gttgttttcc
```

```
5401 ttattaatat tttaaaattt tatacatttt aacagttaca tataaaatag cttactcgac
5461 gacaagactt agcttttaaa agataaaatt ataattagtt ggaataatat taaacgaaac
5521 aaaacaaaag tccaaattcc tcaaacttat agtatttgtt aactttgaaa gaaaaaaagc
5581 atgtttcttt ttttatagtt ttaacttttt ttataaattt acgtaaggtc attaattaaa
5641 atgcgactgt tgtaacggtt ccaaatccaa tgccatttct ggcctcaaca atatgaaatc
5701 taaaagcttc ttcatgagtt agattatcga agtgatgatg tttcataacc acataaataa
5761 ggaacctttg agtcaattac tttacatggt tcataatgat tgttcatgaa cggatatctg
5821 atcgcataat ctaactcaat tttccgtttt aacagtcatt tcatcactgt ataggatttg
5881 acgctcatgg taaaataata ggtaactata atctggttag tatgtaaaaa aataacagca
5941 aagagcctac gcttgagaat ttaagacctc cgattccttt acttccacaa ttgtggtgtc
6001 atccagcaaa aaccactaca aaaattagta aaaatttcca taaaaaattc aaaaaaaatt
6061 taatatactt gattttggta gtatgtatag gcaatatagt gccctgtgtc gagagttcct
6121 ttatgacaca caactgagta taattgataa tccacatcat cttgattaaa gttgtagcgc
6181 attcttaaaa atgctggata agatacttgc ttatctattt tagttgacat ggcaaagtta
6241 ttttgttcaa atcgtttcaa ttgcatgcaa atagttggag gtagtttgtc aaagaccaat
6301 tgtttaatag catttttttga cccgcaacta tgacagctat attgaacttt ctctttggaa
6361 acaaacctaa aataatattt gttaaaaaat ctaaaaacga gcaaatttta cctttccaaa
6421 catccttgaa gtgttggctc attgatgtct aaactaatgt ccatcaaagg atccacagcc
6481 acccgctcct tcttgcaatc taggcatgta acaacgtttt tcagagatcc actaaaaata
6541 cgatgaatag gacaagtgca cggcattgat gttcctcctc cgctttccgt atgcatttgg
6601 tccagcaagt atacaaaaaa ttcgtggcca tcttgctggg agtaaccaca taaagatttg
6661 ctgagtttcc acataagatt taagacagca gtgggaccgt aaaaagtaga tttgttttta
6721 gaattataaa tagatgaaaa catatcgtca atggcacaag tcatgcaagt aggtcgctta
6781 caatcagtag aagtatgaaa cccactgaaa aacaagtttc gaaccaatgg attatgcaaa
6841 atactttgta aaatgacgct catgaaacaa gtagcaccca agttctggat ccccctcaag
6901 cctgctgttg cgcaaactgg agggtatttt cgataggctt ctaagcaaac catttgattg
6961 tacttctcgg gaagtctctt gtgatcactt tgccatgcct tgatattctt gatcttaaaa
7021 cgaagggttt ccaactcggt ctgataaacg tagtcttgac agccaaagca gtaggtatgg
7081 ccattttttta catccacgcc taagaattag taattggtat tcgggaattg catcctaact
7141 cgatattcgc aaaaaggttt cgttgctaaa tcaatagcta tgttttcaac taaacgtacc
7201 gatcatatga gtgtgctcca ttgcgtgttc ttcaccatga tgtccccaaa gacatcccac
7261 agaatgacaa ctcaagcatc gaatactccg cttgttgatg cgtttgcaag tactgcagcg
7321 acgtgggaca aactaaatga aagaaatgtt aaaaggaatc gaatggaaaa agaaaattga
7381 actgtatgaa gtgttgttca caagccttac atgacaactg aatatctggg tacagatctt
7441 ttgatagttt tcaacatccg ccggtttaag tttaagatgc tggcaacctt ctacatcccc
7501 aggcatcctt gatacgtaca aaattggcct ttcagactaa aaagcatctg cttcgatgca
7561 gggtgaaacg acaacaatag acgaactcgc tattctggca aaagaacatt atttctcttt
7621 gcttaaattt gggcaattaa ggatgcggtg tacagcgtta agagccataa agtactcata
7681 gtgtctgcac tactagatac cgatttcgta taaagagcta tcatagctac gctaagtcaa
7741 ttgataaaac accaaaatag atgttcttcg tagttcttga gtgtccttgg ttttcctctt
7801 taaaacataa acttggtaaa ttcccataga tttggctgtt gagtcactgc aaaataaaaa
7861 aaattactgc tagtccttta atcacggttg aatagatgaa attgttaatt ttttaaaagc
7921 ttggtcaaat tagactaatg cctctattaa cgtcgaacgt gtttaagtgt acgcttatgt
7981 ttgttgccat gaaacgtttg cagggctaaa tataaaacca tatatgacta tttaccgtga
8041 gcatctttgt aatacaaccg agtcactgca agcaggtaca tgtacgaaaa tttttgtcta
8101 ttacatttgt atagaattta ttacttaaca attgactagg ggcagatttt tgtgccgaaa
8161 cagtttgtgt taccctaaat agagattgta gggcaacaat acgtctgcag aagttcaaat
8221 atatatgtag caaatgagtg tgacttactt aaaaattggt ttttgatcaa gaaataaaat
8281 cctttggtgc tttttctttt ttatgtctag ccttgttgaa tgtatttttc catcacaatg
8341 ttgatcttgt ctactcaata acacactaga ctatgttaag tacacttgcc cttgcccaat
8401 ctgcaatcgg acatcgtatt ttatttttgc atactgaatt tacgtgtatt cgtagtttca
8461 cgagttaacg attttcgaat atccaactct cattgaactt ataataatcc gtaatccttt
8521 accttttttt tgattacaag tgtctaatta agtttctaa atcaagagaa atcaaagata
8581 caagcaaagc tttttttctt acgtacaatg aagggtttag gtctaagtac gagaaccttt
8641 aattggagct cactttcttc cattctttta ccaagaattc ctttagctac tacaaaagcc
8701 gattccttga ttttggctgt tcgtcacgat aaacaagttt tttctgaaga ttatcgacaa
8761 gttgttgacc agtattttga aacatctccc aaaaaaaatg acattcgtct cttttggaac
8821 actcagggtt ttgttcgtct tgccatagtc cagcttgaag aaaacgtttc tgaaaagtcc
8881 gttcgttcag ctgctgcaga ggcagctaag attttaaagt cgaatggtgc aaagtcgata
8941 gctgtcgatg gtatgggttt tcccaaggat gctgccttgg gtgctgcctt ggccacgtac
```

```
 9001 gatttttctc ttcgtagaga ccatttgagc gtctatcaag atgaaaaggt ggttgaaaag
 9061 gaaaatttat ttacctctcc tgctcctgaa aggttaactt tccaattgct tagcaacaca
 9121 tctgagaaga aaactgcaac cgcagaggag aatgctttta aggtaggttt gattgaagca
 9181 gctgctcaaa acttggctcg ttctcttatg gaatgtcctg ctaattacat gacttctctt
 9241 caattttgtc attttgctca agagttattc caaaattcct caaaggttaa agtattcgtt
 9301 cacgatgaga agtggattga tgagcaaaag atgaacggtt tacttaccgt gaatgctggt
 9361 tccgatattc cacctcgttt cttagaagta caatacattg gtaaagaaaa atcaaaagat
 9421 gatggatggc ttggtttagt tggaaaaggt gtaacgtttg acagcggtgg tatcagtatc
 9481 aagccttctc aaaacatgaa ggagatgcgt gctgatatgg gtggtgctgc tgttatgctt
 9541 tcctctattt atgccttgga acaactttcc atccccgtga atgccgtgtt tgttacccct
 9601 ttaactgaga atcttccttc tggctctgct gctaagccag gagatgttat ttttatgcgc
 9661 aacggtctta gtgtagaaat tgataacact gacgctgagg gtcgtcttat tcttgccgat
 9721 gcagttcact atgtctcttc tcagtataag accaaggctg tgattgaagc cagcacatta
 9781 actggtgcta tgctagtagc cctcggtaat gtatttactg gtgcatttgt tcaaggagaa
 9841 gagttgtgga agaatcttga aactgcttcc catgatgctg gagacctttt ctggcgtatg
 9901 ccatttcacg aggcttacct caagcaactt acaagttctt caaatgctga tttgtgcaat
 9961 gttagtcgtg ccggtggtgg ttgttgtaca gcggctgcct ttattaaatg cttcttagcc
10021 cagaaagatt tgtcttttgc tcatttagac attgccggag ttatggataa gcagcttaat
10081 tcttgggatt gtgacggcat gagcggacgc ccagttcgca ctattattga agtcgctcgc
10141 aagtattagc ttatttattt tcttggctat atgtgcaaac tttatgtttt tacttattta
10201 ctgttgatta caccaatctc atttttcaat gtccttgtta agtgtatgtc aatgatatgt
10261 agctatcatg attttattct ttttaatact gttttccaca accgttcata acggttgtaa
10321 ggaccgagtc tttctgatgg ctctggttct gtttccttta aaaataaaat caattcggtt
10381 gaagtctcat gttgttaaat ttcactttcg aagtagcgaa tttttttttt ctattttgaa
10441 tatagatttg tttattaaat gtagtaatca atttactcat gactattact tttgtcaaac
10501 aaaatcataa tacaaaacta aatctttaaa ggtagaatat tctcctgtgc gagtttacaa
10561 tactttggga tatgatcaaa tcaatttttt ttagatatct gttgaaattt ttataacagt
10621 aattttgtat tatttatcat actatttatt agcgttacct gccctcttaa atgtacacat
10681 catgtggggg tgggggaacc agtgctggag atcacttctt gctataccgt ttggctcatg
10741 gcaactattg aggatagcaa ggagcgttac aatttgcaaa tcaatatgat aaacagaaaa
10801 acaacaagtt gaagcttaac actgtagcaa tataacaatg aaaaaaataa tcagaattaa
10861 agcaagaagc ctataaatgg tccatcaatt gtgaaatcga tattgctaga tgtcaggcaa
10921 atcgtgaact cggttgaaaa atgtcgttaa tattcatgtg aaatcataat tacagcaata
10981 tgttcaaaga cttttggata ctttataata atgaggattc tatttttgag tggtcacatg
11041 gtctcgtaac catgaggtac gccttacttt tacggacgaa acaatatttc acattcagct
11101 ctcaatgtta agtatccaaa caaacatggt acaaaccaat aaagcatgta aattaataag
11161 cttgaaataa ggaatagtca ttaaattcag tttagacata atgtaaaagg tatcaatttt
11221 gtctttggta tagatacaat taagaatttc gagacggaac gttcggtgat gtagagtttt
11281 cataataatt tctgcagttg cttacactaa ccgaagaatc gtccagtata aactgctgac
11341 gatgccttaa atcaataatt gcaaattgca tagttgacgt tttcccgggt ggcacatagt
11401 gagatttgag tcgaataatg aaatagcctc tgacgagtgg tttcgtacca agagtttcaa
11461 aaagaataat ccacattaac tcaatgtatc attgatgacg gaatgttgct tatatagcca
11521 gagtaacccт gctctgaagc atctaaggtt atatcttttt aacttttagc tagcaaactc
11581 acttgttcct ttaagttcag aaagtatctg ctggttattc aacagataat tatcttgcga
11641 aaaaggaata cacactcagc attgtaattt gttgagtata tttcacaagc acattaaatt
11701 gcattatcaa acatgagtgg ggatatttta gtcggtgaat atctattcaa aaggcttgaa
11761 caattagggg tcaagtccat tcttggtgtt ccaggagatt tcaatttagc tctacttgac
11821 ttaattgaga aagttggaga tgagaaattt cgttgggttg gcaataccaa tgagttgaat
11881 ggtgcttatg ccgctgatgg ttatgctcgt gttaatggtc tttcagccat tgttacaacg
11941 ttcggcgtgg gagagctttc cgctattaat ggagtggcag gttcttatgc ggagcatgtc
12001 ccagtagttc atattgttgg aatgccttcc acaaaggtgc aagatactgg agctttgctt
12061 catcatactt taggagatgg agactttcgc actttcatgg atatgtttaa gaaagtttct
12121 gcctacagta taatgatcga taacggaaac gatgcagctg aaaagatc
//
```

Revised: July 5, 2002.

Oct 3 2002 17:48:23

Journal of General Microbiology (1980), 119, 109–115. Printed in Great Britain

09/754,083

Characterization of *Bacillus stearothermophilus* Plasmid pAB124 and Construction of Deletion Variants

By ALISTAIR H. A. BINGHAM,* CHRIS J. BRUTON† AND TONY ATKINSON

Public Health Laboratory Service, Centre for Applied Microbiology and Research, Porton, Salisbury, Wiltshire SP4 0JG

(*Received* 18 *October* 1979; *revised* 26 *January* 1980)

A restriction endonuclease cleavage map of the tetracycline resistance plasmid pAB124, originally isolated from *Bacillus stearothermophilus*, was constructed using ten enzymes. Tetracycline resistance was associated with a 1·95 megadalton (Md) region of pAB124 lying between two *Eco*RI sites, and this region was circularized to produce a viable tetracycline resistance plasmid (pAB224), with two *Eco*RI fragments of pAB124 deleted amounting to 0·95 Md. A second plasmid (pAB524) with one *Eco*RI fragment (0·6 Md) of pAB124 deleted was also constructed. Restriction endonuclease cleavage maps of pAB224 and pAB524 were constructed.

INTRODUCTION

Several plasmids have been isolated from Gram-positive spore-forming bacilli, e.g. *Bacillus pumilus* (Lovett & Bramucci, 1975), *B. megaterium* (Carlton & Helinski, 1969) and *B. subtilis* (Tanaka *et al.*, 1977; Le Hegarat & Anagnostopoulos, 1977), although none of these has been shown to be associated with any phenotypic trait. More recently, a plasmid from *B. cereus* coding for tetracycline resistance and one from *B. subtilis* producing a bacteriocin have been isolated (Bernhard *et al.*, 1978).

The presence of a plasmid in a bacteriocin (thermocin)-producing strain of *B. stearothermophilus* has recently been reported (Sharp *et al.*, 1979), and we have described the isolation of four plasmids from antibiotic-resistant thermophilic bacilli (Bingham *et al.*, 1979), one of which (pAB124) was shown to confer tetracycline resistance on *B. subtilis*. This paper describes the characterization of pAB124 and the construction of deletion variants of this plasmid.

METHODS

Bacterial strains. Thermophilic *Bacillus* strain TB124, isolated as previously described (Bingham *et al.*, 1979), was classified as *Bacillus stearothermophilus* using standard identification procedures (Cowan & Steel, 1974). *Bacillus subtilis* harbouring plasmid pAB124 was prepared by direct plasmid transformation. *Bacillus subtilis* IG20 (*hsr hsm trp*) was kindly provided by A. Docherty, Department of Bacteriology, Bristol University.

Culture media. For plasmid isolation all strains were grown on TYS medium (Bingham *et al.*, 1979) supplemented with tetracycline at 25 $\mu g\ ml^{-1}$. The following minimal medium (SMS) was used for *B. subtilis* ($g\ l^{-1}$): $(NH_4)_2SO_4$, 2; K_2HPO_4, 14; KH_2PO_4, 6; sodium citrate, 1; $MgSO_4.7H_2O$, 0·2; glucose, 5.

Isolation of plasmid DNA. This was carried out with 400 ml cultures using the procedure previously described (Bingham *et al.*, 1979). Plasmid pAB124 was originally isolated from *B. stearothermophilus* TB124 and transformed into *B. subtilis* (Bingham *et al.*, 1979). All experiments described in this paper used pAB124 isolated from *B. subtilis*. Plasmid pUB1654 was a gift from A. Docherty.

† Present address: Department of Biochemistry, Imperial College, South Kensington, London SW7 2AZ.

0022-1287/80/0000-9018 $02.00 © 1980 SGM

Small-scale plasmid preparation. This was used for extracting recombinant DNA from large numbers of clones. Cultures (15 ml) in TYS medium containing the selective antibiotic(s) were grown overnight at 37 °C without shaking. Cells were harvested at 10000 *g*, washed once in 5 ml TES buffer (30 mM-Tris/HCl pH 8·0, 5 mM-Na_2EDTA, 50 mM-NaCl) and resuspended in 100 μl TES. EDTA (25 μl, 0·25 M) and lysozyme (25 μl, 15 mg ml^{-1}) were added and the suspension was shaken gently on ice for 10 min. The suspension was then placed in a 37 °C water bath for 5 min and 25 μl sodium dodecyl sulphate (10 %, w/v) was added, to bring about lysis, followed by 50 μl NaCl (5 M); this mixture was left at 0 °C for 2·5 to 3 h. A cleared lysate was obtained by centrifugation at 45000 *g* for 45 min, and this was then extracted twice with chloroform/3-methylbutan-1-ol (24:1, v/v), once with phenol (freshly distilled over N_2) equilibrated in TES buffer and three times with diethyl ether. The plasmid DNA obtained was suitable for restriction endonuclease analysis.

Plasmid transformation. Bacillus subtilis IG20 was grown overnight at 37 °C in 200 ml SMS medium supplemented with 0·5 ml tryptophan (2 mg ml^{-1}) and then diluted with an equal volume of similar medium (prewarmed). The absorbance (540 nm) was monitored at 30 min intervals, and at a value equivalent to maximum competence (30 to 60 min after cessation of exponential growth) 1 ml samples of the culture were added to plasmid DNA (0·5 to 5 μg) in 50 μl polyethylene glycol 1000 (50 %, w/v) and incubated for 1 h at 37 °C with vigorous aeration. Prewarmed TYS medium (5 ml) was then added and incubation was continued for a further hour prior to plating on selective media. Selective concentrations of antibiotics used were 25 μg ml^{-1} for tetracycline and neomycin and 75 μg ml^{-1} for streptomycin. It was not necessary to induce competence by dilution of the culture into a starvation medium (Stacey, 1968).

*Restriction endonucleases. Eco*RI and *Cau*II were purified by procedures developed in this laboratory and digestions were carried out in buffers previously described (Bingham *et al.*, 1979). *Bst*EII, *Hae*III, *Hha*I, *Hin*dII, *Hpa*I, *Hpa*II, *Tha*I, *Xba*I and T4-DNA ligase were purchased from Uniscience Ltd, Cambridge, and digestions carried out in the buffers recommended by the manufacturer. Reaction mixtures contained 0·5 to 1 μg plasmid DNA in a final volume of 20 to 25 μl, and reactions were terminated by adding Na_2EDTA to a final concentration of 10 mM (*Cau*II, *Bst*EII, *Tha*I) or heating at 66 °C for 10 min. For double digestions involving enzymes with different buffer requirements, plasmid DNA was first digested with the enzyme requiring the buffer of lower ionic strength, and then the buffer was adjusted with 10× concentrated components prior to the addition of the second enzyme.

Agarose gels (0·8 %, w/v) were used as previously described (Bingham *et al.*, 1979), with an *Eco*RI plus *Hin*dIII λ C*I*857 digest providing fragments of standard sizes: 13·4, 3·35, 3·2, 2·8, 2·32, 1·39, 1·27, 1·05, 0·89, 0·59, 0·47 and 0·31 megadaltons (Md) (Murray & Murray, 1975). The sizes of unknown fragments were determined graphically from a 10× enlargement of a photograph (35 mm) and were taken as an average of three estimations.

Cloning of pAB124 *Eco*RI *fragments.* Plasmids pAB124 (1 μg) and pUB1654 (1 μg) were digested to completion with *Eco*RI and then the enzyme was denatured by heating at 66 °C for 10 min. The fragments were precipitated by adding 3 M-sodium acetate to a final concentration of 0·3 M and 2 vol. cold (−20 °C) ethanol. After 10 min in a solid CO_2/ethanol bath the precipitate was collected by centrifugation. After removal of all the ethanol the precipitate was resuspended in T4 ligase buffer (66 mM-Tris/HCl pH 7·6, 6·6 mM-$MgCl_2$, 10 mM-dithiothreitol, 0·4 mM-ATP) and incubated overnight at 4 °C with 0·05 units T4-DNA ligase.

RESULTS

Restriction endonuclease site mapping of pAB124

The fragment sizes obtained by digestion of pAB124 DNA (isolated from *B. subtilis*) with *Bst*EII, *Cau*II, *Eco*RI, *Hae*III, *Hin*dII, *Hha*I, *Hpa*I, *Hpa*II, *Tha*I and *Xba*I are given in Table 1. These sizes were determined graphically using a lambda DNA digested with *Hin*dIII plus *Eco*RI as standards subjected to electrophoresis on the same gels as the unknown fragments. The average value for the total size of pAB124 was close to 2·9 Md. A series of double digestion experiments was carried out in order to construct a cleavage map of pAB124 (Table 2).

The positions of the four single sites (*Bst*EII, *Cau*II, *Hpa*I, *Xba*I) relative to each other were determined taking *Cau*II as the reference point. An *Hpa*I plus *Cau*II digest and a *Bst*EII plus *Cau*II digest both gave a small fragment of 0·35 Md; thus the *Bst*EII and *Hpa*I sites are the same distance from *Cau*II site either on the same side, i.e. very close together, or on opposite sides 0·7 Md apart. The latter position was confirmed by a *Bst*EII plus *Hpa*I double digestion. The *Cau*II plus *Xba*I digestion generated fragments of 1·85 and 1·05 Md. It was assumed that the *Xba*I site was 1·85 Md to the right of *Cau*II and the *Hpa*I

Table 1. *Fragments produced by restriction endonuclease digestion of* pAB124

Digestions and agarose-gel electrophoresis were carried out as described in Methods using plasmid DNA isolated from *B. subtilis.*

Enzyme	Fragment size (Md) A	B	C	Sum of fragment sizes (Md)
*Bst*EII	2·90			2·90
*Cau*II	2·90			2·90
*Hpa*I	2·90			2·90
*Xba*I	2·90			2·90
*Hha*I	2·23	0·70		2·93
*Hind*II	2·30	0·62		2·92
*Hae*III	1·50	1·18	0·25	2·93
*Hpa*II	1·80	0·70	0·42	2·92
*Tha*I	1·84	0·65	0·40	2·89
*Eco*RI	1·95	0·61	0·33	2·89

Table 2. *Fragments produced by digestion of* pAB124 *with two restriction endonucleases*

Double digestions and agarose-gel electrophoresis were carried out as described in Methods using plasmid DNA isolated from *B. subtilis.*

Enzyme pair	Fragment size (Md) A	B	C	D	Sum of fragment sizes (Md)
*Bst*EII + *Cau*II	2·55	0·35			2·90
*Bst*EII + *Hpa*I	2·23	0·70			2·93
*Bst*EII + *Xba*I	2·20	0·70			2·90
*Cau*II + *Hpa*I	2·57	0·35			2·92
*Cau*II + *Xba*I	1·85	1·05			2·90
*Hpa*I + *Xba*I	1·48	1·40			2·88
*Hha*I + *Cau*II	2·20	0·44	0·25		2·89
*Hha*I + *Xba*I	1·59	0·71	0·60		2·90
*Hha*I + *Hpa*I	2·10	0·70	0·10		2·90
*Hind*II + *Cau*II	2·30	0·33	0·28		2·91
*Hind*II + *Xba*I	1·50	0·78	0·62		2·90
*Hind*II + *Hpa*I	2·30	0·61	—		2·91
*Hae*III + *Cau*II	1·24	1·18	0·25	0·24	2·91
*Hae*III + *Bst*EII	1·18	0·89	0·59	0·25	2·91
*Hae*III + *Xba*I	1·30	1·18	0·25	0·18	2·91
*Hae*III + *Hpa*I	1·51	1·19	0·16	0·10	2·96
*Hpa*II + *Cau*II	1·80	0·70	0·42	—	2·92
*Hpa*II + *Bst*EII	1·80	0·42	0·35	0·32	2·89
*Hpa*II + *Xba*I	1·80	0·70	0·38	(0·05)†	2·93
*Hpa*II + *Hpa*I	1·45	0·72	0·42	0·35	2·94
*Tha*I + *Cau*II	1·80	0·42	0·40	0·23	2·85
*Tha*I + *Bst*EII	1·83	0·62	0·40	(0·05)†	2·90
*Tha*I + *Xba*I	1·60	0·65	0·40	0·26	2·91
*Tha*I + *Hpa*I	1·75	0·65	0·40	0·10	2·90
*Eco*RI + *Cau*II	1·55	0·60	0·40	0·33	2·88
*Eco*RI + *Bst*EII	1·90	0·62	0·33	(0·05)†	2·90
*Eco*RI + *Xba*I	1·95	0·33*	0·29	—	2·88
*Eco*RI + *Hpa*I	1·20	0·75	0·60	0·32	2·87

* This fragment was overabundant and assumed to be a 'double'.
† Fragment not detected, but its presence can be predicted.

and *Bst*EII sites were positioned using this orientation. If *Hpa*I lies 0·35 Md to the right of the *Cau*II site a double digestion with *Hpa*I plus *Xba*I would generate 1·48 and 1·40 Md fragments, whereas if it lies 0·35 Md to the left, 0·8 and 2·1 Md fragments would be generated. Since fragments of the former sizes were obtained, the *Hpa*I site must lie to the right of the

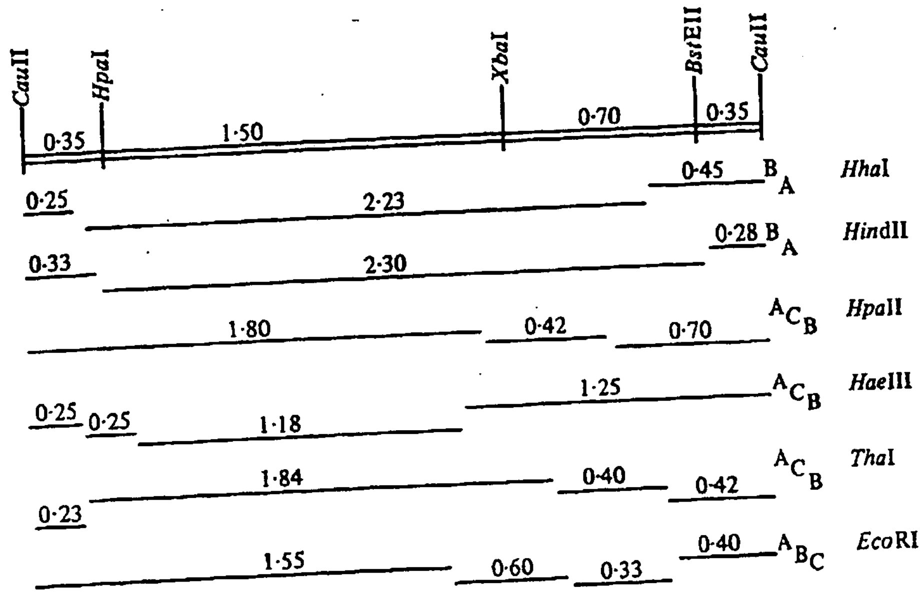


Fig. 1. Relative positions of fragments produced by digestion of pAB124 with several restriction endonucleases.

*Cau*II site with the *Bst*EII site 0·35 Md to the left. The positions of the four sites (Fig. 1) were consistent with the results of all the double digestions carried out.

Using these four single sites as reference points the fragments generated by *Hha*I, *Hin*dII, *Hpa*II, *Hae*III, *Tha*I and *Eco*RI (Table 1) were ordered using the established techniques of double-digestion mapping. For example, the *Eco*RI fragments were ordered easily as only one position of *Eco*RI fragment A was consistent with it containing the *Bst*EII, *Cau*II and *Hpa*I sites but not the *Xba*I site which is present in *Eco*RI fragment B. A *Cau*II plus *Eco*RI digestion generated 1·55 amd 0·4 Md fragments from *Eco*RI A; the 1·55 Md fragment could only lie to the right of the *Cau*II site with the 0·4 Md fragment to the left (Fig. 1). The *Eco*RI B fragment must lie to the right of *Eco*RI A over the *Xba*I site in the position shown (Fig. 1) since an *Eco*RI plus *Xba*I digest yielded 0·33 and 0·29 Md fragments from *Eco*RI B and the distance from *Eco*RI A to the *Xba*I site is 0·35 Md. The remaining fragment, *Eco*RI C, therefore lies to the right of *Eco*RI B.

The fragments generated by digestion of pAB124 with *Hha*I, *Hin*dII, *Hae*III, *Hpa*II and *Tha*I were ordered in a similar manner and their relative positions are shown in Fig. 1.

*Cloning of Eco*RI *fragments of* pAB124 *carrying the tetracycline resistance gene*

Staphylococcal plasmid pUB1654 conferring streptomycin and neomycin resistance was used. This has a single *Eco*RI site lying within the streptomycin resistance gene (A. Docherty, personal communication). It would therefore be expected that insertion of pAB124 fragments at this site would inactivate streptomycin resistance. Transformation of the ligated fragments generated about 35% of clones that were Nm^r Str^s Tc^r. DNA was extracted from several of these clones and was analysed with *Eco*RI (Table 3). The minimum requirement for tetracycline resistance was shown to be the *Eco*RI fragment A (1·95 Md) of pAB124 present alone in 50% of the recombinant molecules examined.

*Construction of small tetracycline resistance plasmids from cloned Eco*RI *fragments of* pAB124

Plasmid DNA from each type of clone, representing all combinations of pAB124 *Eco*RI fragments (Table 3), was digested to completion with *Eco*RI and then ligated with T4-DNA ligase. The DNA was then transformed into *B. subtilis* IG20 and Tc^r Nm^s Str^s clones were identified. Two types of plasmid were isolated from the clones examined (Table 4): pAB224

Table 3. *Restriction endonuclease analysis of recombinant plasmids of* pAB124 *and* pUB1654

Plasmid DNA was extracted from clones of Nm^r Str^s Tc^r phenotype and digested with *Eco*RI as described in Methods.

Clones identified		*Eco*RI digestion		
No.	Plasmid size (Md)	Fragment size (Md)	pAB124 fragment(s) inserted	Type of recombinant
10	6·0	4·0*	—	I
		1·95	A	
3	6·55	4·0*	—	II
		1·95	A	
		0·60	B	
2	6·30	4·0*	—	III
		1·95	A	
		0·32	C	
4	6·90	4·0*	—	IV
		1·95	A	
		0·60	B	
		0·32	C	

* This fragment corresponds to pUB1654.

Table 4. *Eco*RI *analysis of tetracycline resistance plasmids derived from* pAB124:pUB1654 *recombinant molecules*

Plasmid DNA was extracted from Tc^r Nm^s Str^s clones and digested with *Eco*RI as described in Methods.

DNA source*		Tetracycline resistance plasmids			
Type of recombinant	pAB124 *Eco*RI fragment(s) present	No. of clones†	Size (Md)	*Eco*RI fragments (Md)	pAB124 *Eco*RI fragment(s) present
I	A	10	1·95	1·95	A (pAB224)
II	A+B	10	1·95	1·95	A (pAB224)
III	A+C	7	1·95	1·95	A (pAB224)
		3	2·3	1·95, 0·32	A+C (pAB524)
IV	A+B+C	8	1·95	1·95	A (pAB224)
		2	2·3	1·95, 0·32	A+C (pAB524)

* Recombinant plasmids of pAB124 and pUB1654 (see Table 3).
† Ten clones from each transformation were examined.

which was the recircularized *Eco*RI fragment A of pAB124 (1·95 Md) and pAB524 containing *Eco*RI fragments A plus C (2·3 Md). No plasmids containing *Eco*RI fragments A plus B or reconstructed pAB124 were detected.

Plasmid pAB224 was shown to contain seven single sites (*Eco*RI, *Bst*EII, *Cau*II, *Hpa*II, *Hha*I, *Tha*I and *Hpa*I). The position of each site was determined by double digestions and was shown to be consistent with pAB224 being the recircularized *Eco*RI fragment A of pAB124 (Fig. 2). Therefore *Eco*RI A contains the tetracycline resistance genes and all the known essential functions of pAB124.

An *Eco*RI digest of pAB524 generated fragments of 1·95 Md and 0·32 Md. The larger fragment contained *Bst*EII, *Cau*II and *Hpa*I sites and double digestions confirmed that it was the *Eco*RI fragment A of pAB124; the smaller fragment contained *Hha*I and *Hpa*II sites and was assumed to be the *Eco*RI fragment C of pAB124. Two orientations of *Eco*RI C within pAB524 were possible and each would give different sized fragments when digested with *Hpa*II or *Hha*I. The observed sizes of the fragments obtained from five different examples of pAB524 indicated that the order of *Eco*RI fragments A and C was the same as that observed in pAB124 (Fig. 2).

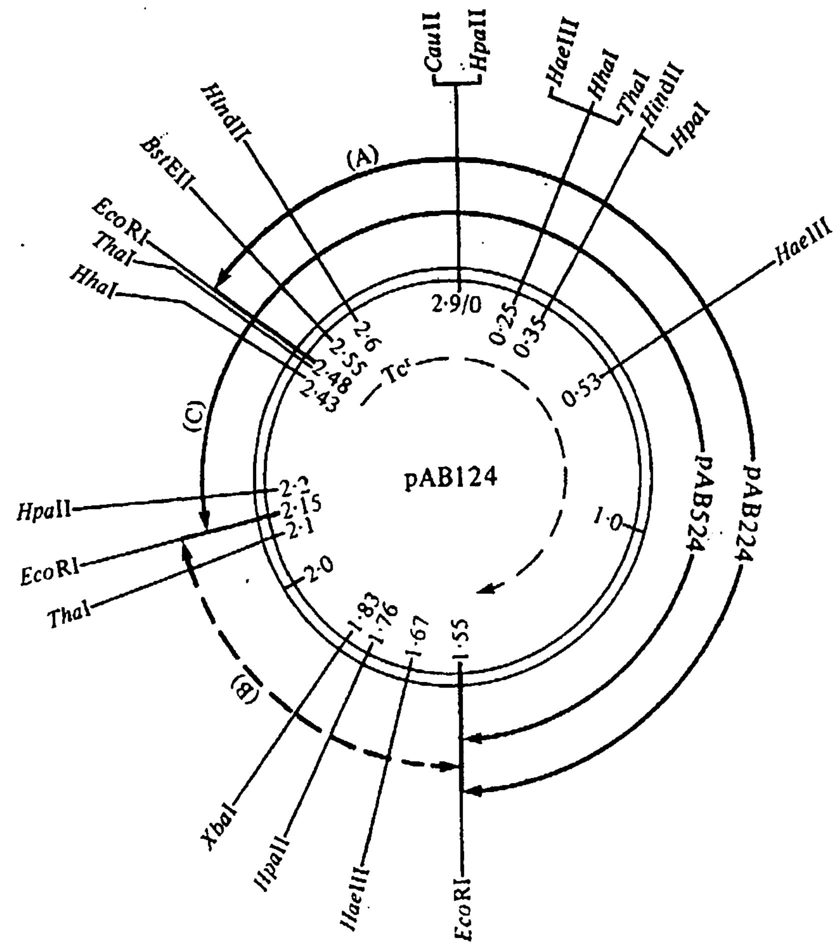
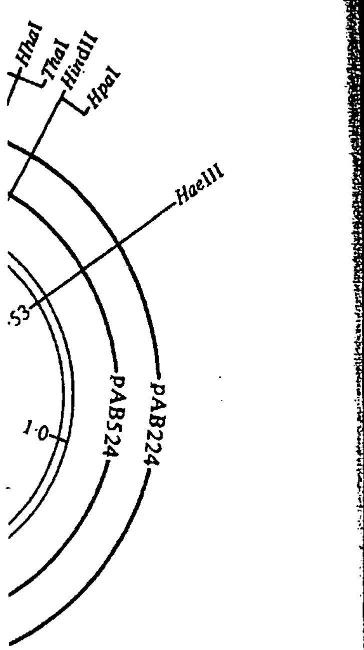


Fig. 2. Restriction endonuclease cleavage map of pAB124, pAB224 and pAB524.

DISCUSSION

Classification of the thermophilic bacilli harbouring pAB124 by standard procedures indicated that the strains were quite closely related to *B. stearothermophilus* but not identical to the neotype strain NCA 1503. The aim of this research is to develop a vector for cloning in *B. stearothermophilus* and pAB124 could provide the basis for such a vector. The expression of pAB124 in *B. subtilis* will also allow the use of this plasmid as a vector for gene cloning in this organism, and it is important to note that the plasmid is stable with no apparent alteration upon uptake and replication in *B. subtilis*, as determined by restriction endonuclease analysis. The transformation frequencies, however, are lower with plasmid DNA isolated from *B. stearothermophilus* than with that from *B. subtilis*; this may in part be due to the formation of concatemers in *B. subtilis* that transform at a higher frequency than the monomeric species (Canosi *et al.*, 1978). The most convenient restriction site for cloning in pAB124 is the single *Xba*I, since this enzyme produces a cohesive terminus and lies within the region of pAB124 not essential for tetracycline resistance or plasmid replication (Fig. 2). We have successfully inserted a staphylococcal neomycin resistance plasmid (pUB110) at the *Xba*I site without inactivating tetracycline resistance. Of the remaining single restriction sites in pAB124, it is not known whether *Bst*EII and *Cau*II produce a cohesive terminus and *Hpa*I is known to produce base-paired termini, thus making the use of these sites for cloning more difficult. In addition, these latter sites lie within the region of pAB124 that appears to be associated with tetracycline resistance.

The development of the staphylococcal vector pUB1654 (Nm^r Str^r; A. Docherty, personal communication) allowed us to examine which region of pAB124 contained the



124, pAB224 and pAB524.

pAB124 by standard procedures
:earothermophilus but not identical
ı is to develop a vector for cloning
asis for such a vector. The expres-
this plasmid as a vector for gene
hat the plasmid is stable with no
btilis, as determined by restriction
however, are lower with plasmid
rom *B. subtilis*; this may in part be
ınsform at a higher frequency than
ıvenient restriction site for cloning
a cohesive terminus and lies within
tance or plasmid replication (Fig.
ıycin resistance plasmid (pUB110)
ce. Of the remaining single restric-
*Cau*II produce a cohesive terminus
s making the use of these sites for
within the region of pAB124 that

654 (Nm^r Str^r; A. Docherty, per-
region of pAB124 contained the

tetracycline resistance genes. Only the large *Eco*RI fragment (A) was essential for the expression of the Tc^r phenotype and the survival of the plasmid as a replicon. The circularized *Eco*RI fragment A (pAB224) contained one site for each of seven restriction endonucleases (Fig. 2); three of these enzymes produce a cohesive terminus (*Eco*RI, *Hpa*II and *Hha*I), and therefore this plasmid is of potential use as a small (1·95 Md) vector for gene cloning. We have attempted to insert at the *Eco*RI site the *B. licheniformis* β-lactamase that is cloned between two *Eco*RI sites in a lambda vector (W.J. Brammar, personal communication) and also the β-lactamase of the staphylococcal plasmid pI258 (Novick *et al.*, 1979). In both cases we could not isolate recombinant molecules containing a β-lactamase gene. We are therefore continuing to examine the ability of pAB224 to express genes cloned at the *Eco*RI site.

The construction of pAB524 in a similar manner to pAB224 gives rise to a plasmid with two *Eco*RI, two *Hpa*II and two *Hha*I sites, but no *Xba*I site. A potentially useful deletion plasmid would be that formed by circularization of *Eco*RI fragments A plus B of pAB124, which would retain the single *Xba*I site and also contain one *Hha*I site. However, repeated ligations of *Eco*RI-digested recombinants of pUB1654 containing *Eco*RI fragments A plus B failed to generate this plasmid; only pAB224 (*Eco*RI A) and pUB1654 containing *Eco*RI fragment A of pAB124 were obtained. This was rather surprising since one would expect circularization of *Eco*RI fragments A plus B to be about as efficient as circularization of *Eco*RI fragments A plus C (pAB524; Table 4).

We intend to examine the ability of pAB124 and pAB224 to express chromosomal functions cloned at the *Xba*I and *Eco*RI sites, respectively, in order to evaluate their use as vectors for gene cloning in *Bacillus* species.

We thank A. Docherty for providing pUB1654 DNA. This work was supported by a grant from the Health and Safety Executive.

REFERENCES

BERNHARD, K., SCHREMPF, H. & GOEBEL, W. (1978). Bacteriocin and antibiotic resistant plasmids in *Bacillus cereus* and *Bacillus subtilis*. *Journal of Bacteriology* 133, 897–903.

BINGHAM, A. H. A., BRUTON, C. J. & ATKINSON, T. (1979). Isolation and partial characterization of four plasmids from antibiotic-resistant thermophilic bacilli. *Journal of General Microbiology* 114, 401–408.

CANOSI, U., MORELLI, G. & TRAUTNER, T. A. (1978). The relationship between molecular structure and transformation efficiency of some *S. aureus* plasmids isolated from *B. subtilis*. *Molecular and General Genetics* 166, 259–267.

CARLTON, B. C. & HELINSKI, D. R. (1969). Heterogeneous circular DNA elements in vegetative cultures of *Bacillus megaterium*. *Proceedings of the National Academy of Sciences of the United States of America* 65, 592–599.

COWAN, S. T. & STEEL, K. J. (1974). *Identification of Medical Bacteria*, 2nd edn. Cambridge: Cambridge University Press.

LE HEGARAT, J.-C. & ANAGNOSTOPOULOS, C. (1977). Detection and characterisation of naturally occurring plasmids in *Bacillus subtilis*. *Molecular and General Genetics* 157, 167–174.

LOVETT, P. S. & BRAMUCCI, M. G. (1975). Plasmid deoxyribonucleic acid in *Bacillus pumilus* and *Bacillus subtilis*. *Journal of Bacteriology* 124, 484–490.

MURRAY, K. & MURRAY, N. E. (1975). Phage lambda receptor chromosome for DNA fragments with *Hind*III and *Eco*RI. *Journal of Molecular Biology* 98, 551–564.

NOVICK, R. P., MURPHY, E., GRYCZAN, T. J., BARON, E. & EDELMAN, I. (1979). Penicillinase plasmids of *Staphylococcus aureus*: restriction-deletion maps. *Plasmid* 2, 109–129.

SHARP, R. J., BINGHAM, A. H. A., COMER, M. J. & ATKINSON, A. (1979). Partial characterization of a bacteriocin (thermocin) from *Bacillus stearothermophilus* RS93. *Journal of General Microbiology* 111, 449–451.

STACEY, K. A. (1968). Synthesis and properties of DNA: transformation. In *Experiments in Microbial Genetics*, pp. 46–51. Edited by R. C. Clowes & W. Hayes. Oxford: Blackwell Scientific Publications.

TANAKA, T., KURODA, M. & SAKAGUCHI, K. (1977). Isolation and characterisation of four plasmids from *Bacillus subtilis*. *Journal of Bacteriology* 129, 1487–1494.

1: L09136. pUC18c cloning ve...[gi:209210]

Links

```
LOCUS       SYNPUC18CV              2686 bp    DNA     circular SYN 27-APR-1993
DEFINITION  pUC18c cloning vector (beta-galactosidase mRNA on complementary
            strand).
ACCESSION   L09136
VERSION     L09136.1  GI:209210
KEYWORDS    beta-galactosidase; cloning vector.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 2686)
  AUTHORS   Gilbert,W.
  TITLE     Obtained from VecBase 3.0
  JOURNAL   Unpublished (1991)
COMMENT     Original source text: Synthetic construct DNA.
            These data and their annotation were supplied to GenBank by Will
            Gilbert under the auspices of the GenBank Currator Program. pUC18c
            - Cloning vector (beta-galactosidase mRNA on complementary strand)
            ENTRY PUC18C                    #TYPE DNA CIRCULAR TITLE pUC18c -
            Cloning vector
                          (beta-galactosidase mRNA on complementary strand)
            DATE      02-APR-1986
               #sequence  16-DEC-1986
            ACCESSION VB0032
            SOURCE    artificial
            COLLECTION  ATCC 37253
            REFERENCE
               #number  1
               #authors  Norrander J., Kempe T., Messing J.
               #journal Gene (1983) 26: 101-106
            REFERENCE
               #number 2
               #authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.
               #book    Cloning Vectors, Elsvier 1985 and supplements
               #comment  vector I-A-iv-20
            COMMENT
               Assembled from pUC19c and M13mp18 by F. Pfeiffer, MPI,
            Martinsried
               Revised 16-DEC-1986 by F. Pfeiffer:
               1062/3 'AT' to 'TA' to match revised sequence of PBR322
               The beta-galactosidase mRNA sequence including the multiple
               cloning site of M13mp18 is on the strand complementary to that
               shown.
            KEYWORDS
            CROSSREFERENCE
               #complement
                 VecBase(3):pUC18
            PARENT
```

```
               Features of pUC18c (2686 bp)
                 residue     source
                   1- 137   2074-2210 pBR322
                 138- 237   2252-2351 pBR322
                 238- 396   1461-1303 (c) Lac-Operon
                 399- 455     57-   1 (c) polylinker of M13mp18
                 456- 682   1297-1071 (c) Lac-Operon
                 683-2686   2352-4355 pBR322
               Conflict (cfl) and Mutations (mut):
                    pUC18c  source
               mut 1942  A  G  3611   pBR322
               mut 2243  T  C  3912   pBR322
            FEATURE
               1629-2417 789-1 (c) Ap-R; b-lactamase
            POLYLINKER  EcoRI-SacI-KpnI-SmaI-BamHI-XbaI-SalI-PstI-SphI-HindIII
            SELECTION
               #resistance  Ap
               #indicator beta-galactosidase
            SUMMARY  pUC18c    #length 2686   #checksum 3662.
FEATURES             Location/Qualifiers
     source          1..2686
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT      666 a    677 c    684 g    659 t
ORIGIN
        1 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca
       61 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg
      121 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc
      181 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc
      241 attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat
      301 tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt
      361 tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttgcatgc ctgcaggtcg
      421 actctagagg atccccgggt accgagctcg aattcgtaat catggtcata gctgtttcct
      481 gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt
      541 aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc
      601 gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg
      661 agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg
      721 gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca
      781 gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac
      841 cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gcccccctga cgagcatcac
      901 aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg
      961 tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac
     1021 ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat
     1081 ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag
     1141 cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac
     1201 ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt
     1261 gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttggt
     1321 atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc
     1381 aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga
     1441 aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac
     1501 gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc
     1561 cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct
     1621 gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca
     1681 tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct
     1741 ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca
     1801 ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc
     1861 atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg
     1921 cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct
     1981 tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa
     2041 aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta
```

```
2101 tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc
2161 ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg
2221 agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa
2281 gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg
2341 agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc
2401 accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg
2461 gcgacacgga aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat
2521 cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata
2581 ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc
2641 atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc
//
```

Revised: July 5, 2002.

Jan 21 2003 18:08:12

EXHIBIT N

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of
E. GREEN *et al.*
Serial No. 09/754,083
Filed: January 5, 2001

Group Art Unit: 1652
Examiner: M. N. RAO
Atty. Docket No. 000487.00007

For: **ETHANOL PRODUCTION**

DECLARATION

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

We, Edward Green, Muhammad Javed, and Namdar Baghaei-Yazdi, declare that:

1. We are the named inventors of the subject matter claimed in the above-identified application. We deposited Bacillus strains LN (accession number 41038), TN (accession number 41039), LN-S (J8) (accession number 41040), and LN-D (accession number 41041) with the National Collections of Industrial, Food and Marine Bacteria (NCIMB) at 23 St. Machar Drive, Aberdeen AB2 1RY, Scotland, United Kingdom. Each deposit was made under the terms of the Budapest treaty.

2. Upon issuance of a patent each of these strains will be irrevocably and without restriction or conditions be available to the public.

3. We hereby declare that all statements made herein of our own knowledge are true and that all statements made on information and belief are believed to be true; and further that these were made with the knowledge that false statements made willfully are punishable by fine, imprisonment, or both a fine and imprisonment under Section 1001 of Title 18 of the United States; and further that false statements made willfully may jeopardize the validity of any patent issuing on an application in which the false statements were made.

20/01/03
Date

Edward Green

17/01/03
Date

Muhammad Javed

17/1/03
Date

Namdar Baghaei
Namdar Baghaei-Yazdi

oese, C.R. and
–296.
. (1984) Appl.
(1963) J. Biol.
A. and Crabill,
1162–1169.
Appl. Environ.
er, R.K. (1982)
ppl. Microbiol.
ay, W.D. and
l. 48, 437–447.

FEMS Microbiology Letters 26 (1985) 333–336
Published by Elsevier

FEM 02008

Mutants of *Bacillus stearothermophilus* lacking NAD-linked L-lactate dehydrogenase

(Fermentation; ethanol production; fluoropyruvate resistance; thermophile)

Mark A. Payton and Brian S. Hartley *

*Department of Microbiology, Biogen S.A., P.O. Box 1211, Geneva 24, Switzerland, and * Centre for Biotechnology, Imperial College, London SW7 2AZ, U.K.*

Received 15 November 1984
Accepted 23 November 1984

1. SUMMARY

Bacillus stearothermophilus produces low levels of ethanol following fermentation of glucose at temperatures in excess of 60°C. A major by-product is L-lactic acid. To improve ethanol yields mutants of this organism were isolated which lacked NAD-linked L-lactate dehydrogenase, the enzyme responsible for lactic acid production. This was achieved by selection for resistance to a suicide substrate, fluoropyruvate. Such mutants no longer produced lactic acid and yields of ethanol from glucose were increased by at least a factor of two.

2. INTRODUCTION

B. stearothermophilus is a facultatively anaerobic thermophilic bacillus capable of growing either oxidatively or fermentatively on glucose [1]. During fermentative growth in rich broth with glucose or sucrose as the major carbon source *B. stearothermophilus* displays a "mixed-acid" type fermentation with the production of lactic, acetic and formic acids and some ethanol [2].

The oxidation and fermentative metabolism of glucose by wild-type *B. stearothermophilus* (NCA1503) has been described previously [3,4]. It appears similar to that typified by the enteric bacterium *Escherichia coli* in that under aerobic, oxidative conditions, pyruvate formed by glycolysis is further metabolised by a pyruvate dehydrogenase (PDH) complex [5] to form acetyl CoA, which would normally be oxidised through the citric acid cycle. Growth under fermentative conditions, however, induces the synthesis of a pyruvate formate lyase (PFL) system [6] which converts the pyruvate to acetyl CoA, but with the concomitant production not of NADH as with the PDH complex but of formate. It has been reported previously both for *B. stearothermophilus* itself [2] and for other "mixed-acid" type fermentative microorganisms such as *E. coli* [7] that the pH of a fermentation can drastically affect the ratios of the different possible end-products. For *B. stearothermophilus*, there appears to be a relationship whereby the production of low levels of lactic acid as a fermentative end-product facilitates the production of relatively higher amounts of ethanol. This is due, presumably, to the extra availability both of carbon as pyruvate and of reducing equivalents as NADH under such conditions. We have been investigating the effects of a variety of parameters on the overall fermentation balance of

this organism as a model system for the improvement of ethanol-producing thermophiles for possible use in a thermophilic ethanol fermentation process. In order to improve ethanol yields, therefore, we attempted to abolish completely the production of lactic acid during the fermentation of glucose by *B. stearothermophilus* by the manipulation of media and growth conditions, but without total success. We therefore opted to remove NAD-linked lactate dehydrogenase, the enzyme responsible for L-lactate production, by mutagenesis. We describe here the isolation of mutants lacking this enzyme and a preliminary discussion of their fermentation characteristics.

3. MATERIALS AND METHODS

The starting strain of *B. stearothermophilus* used in this study was a prototrophic variant of *B. stearothermophilus* NCA1503. This strain, PSII, was isolated previously as a strain capable of rapid anaerobic growth on glucose (Payton, unpublished). Cultures were grown and maintained on LB glucose medium containing yeast extract (20 $g \cdot l^{-1}$), tryptone (10 $g \cdot l^{-1}$) and D-glucose (2 $g \cdot l^{-1}$) at a final pH of 7.0. For plates, LB glucose medium was solidified with 1.5% (w/v) Difco Bacto Agar. Growth was performed at 60°C unless otherwise stated.

Fluoropyruvic acid (sodium salt) was obtained from Sigma (London). Mutants of *B. stearothermophilus* PSII resistant to this compound were isolated first by growing PSII overnight at 60°C aerobically in LB glucose medium. Samples of 0.1 ml of these cultures containing approximately 10^8 viable cells were then plated onto LB glucose medium solidified with agar and allowed to dry. Fluoropyruvate was then applied, as a solid (2–5 mg) to the centre of each plate and the plates were incubated overnight at 60°C, after which time zones of inhibition of 1–3 cm could be seen within a lawn of bacterial growth. Occasionally single colonies could be seen within the zone of inhibition and these were picked onto LB glucose solid medium for purification. Single colonies purified in this way were retested by repeating the original isolation procedure. Only those strains which exhibited zones of inhibition repeatedly and significantly less than the parental strain PSII (i.e. < 0.5 cm) were chosen for further study.

Initial classification of these strains was performed by growing the strains individually in 2-ml volumes of LB glucose medium in static culture at 60°C for 12 h. Half of the culture was centrifuged at $5000 \times g$ for 5 min and the resultant supernatant assayed for the presence of L-lactic acid using a commercially available kit for the determination of L-lactic acid (Boehringer). The remaining half of the culture was permeabilised by a modification of the method previously published for *E. coli* [8].

To 1 ml samples of culture, 10 μl aliquots of a 10% (v/v) solution of toluene in absolute ethanol were added. Following incubation on ice for 20 min the mixtures were vortexed for 1 min and returned to ice. Samples were used within 30 min of preparation to measure NAD-linked L-lactate dehydrogenase by the method of Tarmy and Kaplan [9].

The assay contained in a total volume of 1 ml: NADH (0.3 mM), sodium pyruvate (30 mM) in potassium phosphate buffer (0.1 M, pH 7.5). Sufficient permeabilised extract was added to give an absorbance change of approx. 0.1 min^{-1} at 340 nm and 25°C. Assay results were verified by preparing sonicates of the appropriate strain; cells from 100 ml overnight cultures in LB glucose medium were harvested by centrifugation, washed twice in 10 ml volumes of potassium phosphate buffer (0.1 M, pH 7.5) and finally resuspended in 5 ml of the same buffer. Suspensions were sonicated for 5 1-min bursts (150 W MSE ultrasonic disintegrator, peak to peak amplitude 12 μm) at 0°C and centrifuged at 4°C and $12000 \times g$ for 15 min to remove debris. Lactate dehydrogenase was then assayed in supernatants as described above.

Determinations of the products of glucose fermentation were performed on supernatants from cultures grown on either CG medium (tryptone 30 $g \cdot l^{-1}$, yeast extract 10 $g \cdot l^{-1}$, D-glucose 10 $g \cdot l^{-1}$) or SG medium (tryptone 1 $g \cdot l^{-1}$, yeast extract 0.5 $\cdot$ $g\ l^{-1}$, D-glucose 10 $g \cdot l^{-1}$) for 40 h at 57°C. Assays for lactic acid were performed as described

ıs which ex-
· and signifi-
ıII (i.e. < 0.5

ns was per-
ıally in 2-ml
tic culture at
; centrifuged
ıltant super-
ʟ-lactic acid
for the de-
ger). The re-
abilised by a
ly published

aliquots of a
əlute ethanol
n ice for 20
· 1 min and
ithin 30 min
ked ʟ-lactate
Tarmy and

ume of 1 ml:
(30 mM) in
H 7.5). Suffi-
d to give an
nin^{-1} at 340
ified by pre-
strain; cells
LB glucose
ition, washed
m phosphate
suspended in
ıs were soni-
SE ultrasonic
le 12 μm) at
00 × g for 15
rogenase was
ibed above.
f glucose fer-
natants from
ı (tryptone 30
ose 10 g · l^{-1})
yeast extract
0 h at 57°C.
l as described

above. Ethanol and acetate were also measured using commercially available kits from Sigma (U.K.) and Boehringer, respectively.

4. RESULTS AND DISCUSSION

We had previously failed to isolate lactate dehydrogenase-deficient strains by conventional techniques such as screening for low acid-producing mutants of *B. stearothermophilus* PSII following growth on glucose either aerobically or anaerobically. We therefore sought an alternative method based on a selection for resistance to a suicide substrate, fluoropyruvate. We had previously noted that the pyruvate analogue fluoropyruvate (FCH_2–CO–COOH) was toxic to *B. stearothermophilus*. The toxicity of fluoropyruvate could depend on its ability itself to inhibit a key metabolic step, or could depend on its conversion initially either to fluoroacetyl CoA or to fluorolactate. In each case, amongst fluoropyruvate-resistant strains, some should be defective in uptake or accumulation of the toxic analogue. In addition, however, mutants should be isolated which affect the target reaction of inhibition or the metabolism of fluoropyruvate to fluoroacetyl CoA (via PDH or PFL) or to fluorolactate (via NAD-linked ʟ-lactate dehydrogenase).

We therefore isolated fluoropyruvate-resistant mutants (MATERIALS AND METHODS) and proceeded to characterise the resulting strains. Strains chosen for further study, designated FPY (fluoropyruvate-resistant) mutants, were analysed by screening glucose-grown cells for the presence or absence of NAD-linked ʟ-lactate dehydrogenase. This was tested initially on cells permeabilised with toluene, then verified using cell extracts prepared by sonication (MATERIALS AND METHODS). The supernatants from the same cultures were also assayed for the presence of ʟ-lactic acid (MATERIALS AND METHODS). The results of these analyses are shown in Table 1. Of 12 strains thus tested 8 failed to produce lactic acid and also lacked detectable LDH activity in permeabilised cell suspensions. Assays performed on sonicates of selected strains were used to verify the absence of LDH. One of these strains, FPY 15, was chosen for further study and was redesignated *Ild*15 (ʟ-lactate dehydrogenase-deficient) and its fermentation balance studied.

Table 1

An analysis of fluoropyruvate-resistant mutants of *B. stearothermophilus* PSII

Strain No.	Lactate production	LDH detected in toluenised cells	LDH present in sonicates
FPY 4	+	+	+
FPY 5	−	−	−
FPY 6	−	−	−
FPY 7	+	+	+
FPY 8	+	+	NT [a]
FPY 9	+	+	NT
FPY 10	−	Low	NT
FPY 11	−	Low	NT
FPY 12	−	−	NT
FPY 13	−	Low	NT
FPY 14	−	−	NT
FPY 15	−	−	−

[a] NT, not tested.

Table 2 shows a comparison of the major fermentation products from glucose of *Ild*15 vs. its parent PSII. First, in both strains acetate formation is higher on complex medium than on simple medium. We believe this is due to conversion of amino acids in the complex medium into acetate. Second, ethanol production by *Ild*15 is higher by a factor of two than by PSII. We have subsequently performed a number of fermentations on *Ild*15 both in batch and continuous culture at varying pH which have been published elsewhere [3,4] and have improved the ethanol yield by *Ild*15. The figure quoted for PSII of 0.5 mol ethanol per

Table 2

Yields of fermentation products from glucose (mol product/mol glucose) for *B. stearothermophilus* strains PSII and *Ild*15

The media used and assays of products are described in MATERIALS AND METHODS.

Strain	PSII		*Ild*15	
Medium	CG	SG	CG	SG
Lactate	1	1	0.04	0
Acetate	2	0.5	3.38	1
Ethanol	0.5	0.5	1	1
Total	3.5	2.0	4.42	2.0

mol glucose is, however, the highest we have obtained to date; at lower pHs ethanol production drops off sharply. These data show first that mutants of *B. stearothermophilus* lacking NAD-linked lactate dehydrogenase retain the ability to grow on glucose aerobically or anaerobically. Furthermore, the lactate dehydrogenase-deficient (*lld*) strains of *B. stearothermophilus* produce significantly more ethanol than the wild-type strain following fermentation of glucose, verifying the trends shown by McKray and Vaughn [2]. Results from fermentations performed on *lld*15 suggest that the isolation of similar mutants in other organisms of potential use in ethanol fermentations may be a means of increasing ethanol yields. In addition, the isolation of *lld* mutants of *B. stearothermophilus* amongst fluoropyruvate-resistant strains suggests that in this organism one mode of fluoropyruvate toxicity is dependent on its initial conversion to fluorolactate, in contrast to previous reports of the effect of this and similar fluoroanalogues on *E. coli* [10]. We have, in fact, shown (Payton, M.A., unpublished) that fluoropyruvate is a substrate, albeit a poor one, for semipurified LDH from *B. stearothermophilus*, as has been reported for LDHs from other sources [11].

REFERENCES

[1] Atkinson, A., Ellwood, D.C., Evans, C.G.T. and Yeo, R.G. (1975) Biotech. Bioeng. 17, 1375–1377.

[2] McKray, G.A. and Vaughn, R.H. (1957) Food Res. 22, 494–500.

[3] Hartley, B.S. and Payton, M.A. (1983) Biochem. Soc. Symp. 48, 133–146.

[4] Hartley, B.S., Payton, M.A., Pyle, D.L., Mistry, P. and Shama, G. (1983) in Biotech 83, pp. 895–905. Online, London, U.K.

[5] Henning, U., Dennert, G., Hertel, R. and Shipp, W.S. (1966) Cold Spring Harbor Symp. Quant. Biol. 31, 227–234.

[6] Knappe, J., Blaschkowski, H.P., Grobner, P. and Schmitt, T. (1974) Eur. J. Biochem. 50, 253–263.

[7] Dawes, E.A. and Foster, S.M. (1956) Biochim. Biophys. Acta 22, 253–265.

[8] Reeves, R.E. and Sols, A. (1973) Biochem. Biophys. Res. Commun. 50, 459–466.

[9] Tarmy, E.M. and Kaplan, N.O. (1968) J. Biol. Chem. 243, 2579–2586.

[10] Avi-Dor, Y. and Mager, J. (1956) J. Biol. Chem. 222, 249–258.

[11] Eisman, E.H., Lee Jr., J.A. and Winer, A.D. (1965) Biochemistry 4, 606–610.